DOCUMENT RESUME

ED 243 246 EA 016 742

AUTHOR	Williams, Richard C.; Bank, Adrianne
TITLE	Improving Instruction through the Management of Testing and Evaluation Activities: A Guidebook for School Districts.
INSTITUTION	California Univ., Los Angeles. Center for the Study of Evaluation.
SPONS AGENCY	National Inst. of Education (ED), Washington, DC.
PUB DATE	Nov 82
GRANT	NIE-G-80-0112,P5
NOTE	299p.; Evaluation Design Project. Supersedes ED 226 016.
PUB TYPE	Guides - Non-Classroom Use (055)
EDRS PRICE	MF01/PC12 Plus Postage.
DESCRIPTORS	Academic Achievement; Case Studies; *Change Strategies; Curriculum Evaluation; *Educational Planning; Educational Strategies; Elementary Secondary Education; *Evaluation Utilization; Instructional Development; *Instructional Improvement; *Program Evaluation; *Test Interpretation; Test Use

ABSTRACT

This guidebook shows how school district personnel from many areas--curriculum, instruction, supervision, staff development--can coordinate their activities to focus on instructional improvement. It reveals how data derived from tests, when properly used, can be useful in helping district personnel work with school and community people to assess the adequacy of an instructional program and to improve it. After an introductory overview of the guide, chapter 2 presents a scenario of a week in a relatively small school district where declining test scores have just been released to the newspapers. The ongoing activities, daily crises, and events of the school day form a backdrop against which board members and administrators attempt to assess the significance of the scores and to plan their actions accordingly. Chapter 3 addresses assumptions underlying questions or demands for action from the public in response to release of test scores. These assumptions wrongly assume that a single factor is responsible for low achievement, but actually no single procedure, material, facility, or person operates in isolation from the complex interactions in an educational system. In chapter 4, six districts tell how they responded in a unique way to the problem of increasing student achievement. Each used test scores and evaluative data not only to monitor but to plan instruction. Chapter 5 identifies five reasons why a long-term coordinated district strategy is difficult to achieve. An open systems perspective then shows how changed assumptions can direct a school district toward a coherent plan to interpret testing, evaluation, and instruction. Chapter 6 provides (1) reasons for developing a unique evaluation system, (2) starting points (opportunities and constraints), (3) a sequence for development, (4) selection of emphasis, and (5) methods for identifying sources of support. (TE)

Center for the Study of Evaluation

UCLA Graduate School of Education
Los Angeles, California 90024

ED243246

Deliverable - November 1982

EVALUATION DESIGN PROJECT

IMPROVING INSTRUCTION THROUGH THE MANAGEMENT OF TESTING AND EVALUATION ACTIVITIES

A GUIDEBOOK FOR SCHOOL DISTRICTS

SCOPE OF INTEREST NOTICE

The ERIC Facility has assigned this document for processing to: EA

TM

In our judgement, this document is also of interest to the clearinghouses noted to the right. Indexing should reflect their special points of view.

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☒ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

- Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

EA 016 742

Deliverable - November 1982

EVALUATION DESIGN PROJECT

IMPROVING INSTRUCTION THROUGH THE MANAGEMENT OF TESTING AND EVALUATION ACTIVITIES

A GUIDEBOOK FOR SCHOOL DISTRICTS

Richard C. Williams and Adrianne Bank
Project Directors

Grant Number
NIE-G-80-0112, P5

CENTER FOR THE STUDY OF EVALUATION
Graduate School of Education
University of California, Los Angeles

The project presented or reported herein was performed pursuant to a grant from the National Institute of Education, Department of Education. However, the opinions expressed herein do not necessarily reflect the position or policy of the National Institute of Education, and no official endorsement by the National Institute of Education should be inferred.

TABLE OF CONTENTS

PREFACE 1

CHAPTER 1: What This Book is About 4
in which the major perspectives of the work are summarized along with a precis for each of the remaining chapters . . .

CHAPTER 2: School District Scenario 9
in which district personnel consider public reaction to test score decline as reported in the media . . .

CHAPTER 3: School District Questions 26
in which four simple questions implying "quick fix" solutions are reframed for more productive thinking . . .

CHAPTER 4: School District Stories 76
in which six districts tell how they developed and implemented systems linking testing, evaluation and instruction to improve student learning . . .

CHAPTER 5: School District Analyses 202
in which the impediments to developing such systems are discussed along with the conditions propelling districts in this direction . . .

CHAPTER 6: School District Possibilitites 233
in which district personnel, who are enroute to their own instructional renewal, examine their own situations in terms of impediments as well as facilitating conditions . . .

PREFACE

From the beginning of formal education, parents and pupils and teachers have been interested in, and concerned about what and how children learn.

In the past, when pupils did not perform at what adults regarded as acceptable levels the blame was often placed on the pupils' lack of ability or unwillingness to apply him or herself to the academic tasks that were set out. Of course, teachers received a share of the blame as well.

More recently, notice has been taken of the role that parent and peer attitudes may play in influencing students to value, and to prosper in an educational system.

Even so, in addressing the issue of how to increase learning, the focus of analysis has been on the unique interchange between teacher and pupil as it takes place within the school classroom.

In the last 20 years, federal and state governments have begun funding specific programs intended to ameliorate what some have seen as persistent problems of the schooling process Riding along with this external funding has come a new focus for educational analysis, namely the extent to which pupil achievement is influenced by educational programs and by school district level actions.

This shift in level of analysis has been accompanied by a new set of questions added to concern with pupil motivation, teacher behavior, peer interactions, and parent encouragement. These concerns with program delivery processes, adequacy of implementation assessed by administrative monitoring, and testing of student achievement, which had historically been a measuring tool so that teachers could determine pupil achievement, became folded into the newer field of program evaluation.

Program evaluation used tests to assess pupil programs and other evaluative techniques -- classroom visitation, surveys, opinionnaires, to describe program inputs. Ideally, data about the educational messages delivered to students and student learning from those messages should lead school personnel to a determination of where action should be taken to enhance student learning.

This book is our attempt to analyze the bridges that a district can build between test and evaluation data on the one hand, and school district instructional decision making on the other. Our analysis is not focused on the teacher in the classroom, nor on the principal in the school building. Neither is it on the administrator of the special program. Rather, it is on the personnel in the school district's central office. We have selected this level because we believe many, but certainly not all, pupil achievement problems can be and, indeed, must be, addressed at that level. Given the way our public school system is presently structured, the ultimate operational responsibility for pupil achievement must rest at the district level.

This book, focusing on the district level, represents a relatively new approach to this problem of improving pupil achievement.

We suppose that if there were any word that would typify the book's spirit and structure, it would be the word synthesis. This synthesis is represented in several ways: the subject matter, the operational level, and the contributors. To elaborate: First, the subject matter represents a synthesis of several fields of inquiry. We attend to the characteristics of and interrelationships between testing, evaluation and instruction on the one hand, and organizational theory on the other. Second, at the operational level, we try to integrate theory and practice, or put another way, analysis with action. There is no formula or pre-fabricated program that will guarantee a school district an effective way to link testing, and evaluation with instruction to improve pupil achievement. Developing such a program depends upon analyzing each district's unique problems, make-up and setting and devising a program to meet local realities. Third, the book's contributors represent experiences gained through detailed observation and academic searchings coupled with experiences gained through daily struggle with designing, developing, and implementing programs for linking testing and evaluation with instruction in school districts.

Blending and synthesizing ideas generated from diverse experiences has been sometimes difficult, but always stimulating. While our own knowledge and interest in this topic has grown dramatically in the process of conducting our research and working with a wide array of individuals, the ultimate success of the book will rest on whether or not our work helps school districts to better develop and implement more effective instructional programs.

CHAPTER 1

What This Book is About

in which the major perspectives of the book are summarized along with a precis for each of the remaining chapters . . .

CHAPTER 1

What This Book is About

This book is about how school districts can help their students learn more.

This book says that school district personnel who work in many areas -- curriculum, instruction, supervision, staff development -- can coordinate their activities so as to focus on instructional improvement.

This book says that data derived from tests, when properly used, analyzed, interpreted, can be useful in helping district personnel in many areas work with school and community people to assess the adequacy of the instructional program and to improve it.

Many districts, at present, collect a lot of data. Much of it remains merely data, undigested, unused, or worse, transformed into misinformation. Low test scores, when reported by the media, and read by the public may appear to be an accurate assessment of the lack of quality of schools and the inadequacy of student achievement. The problem, then, is framed by the question, "How to we raise test scores?"

> Chapter to presents a scenario of a week in a relatively small school district where test scores have just been released to the newspapers. The ongoing activities, daily crises, and events, the personalities of individuals, form the backdrop against which board members, the superintendent, administrators and teachers ask themselves and one another - "What do the scores mean?" "What next?" "Where do we go from here?"

Education has the unique distinction of being a topic that almost every segment of society feels equipped to judge and express convictions. These feelings and convictions stem from the fact that the American system of public education has offered an educational program for all youth, and from the personal experience individuals tend to draw retrospective conclusions regarding the relation of the educational experience to successes, failures, wishes and unsolved problems both during school and in the post-school years.

A common reaction to the many problems of society is to look for a cause or reason for the undesired circumstances. In the search for a single cause or reason for complex problems, it is common for a single content, process, or procedure to be identified as the "cause" and to suggest that a change will eliminate the problem.

The organization, administration and practices of education are under close scrutiny by all segments of society. The questions raised about educational practices are often not asked to obtain explanation or information, but are statements of belief regarding a solution to a perceived problem. For example, questions such as -- "How soon can we expect test scores to go up?" "Why don't we identify incompetent teachers and do something about them?" "Which district has an instructional program that has been proven to improve test scores?" and "Why don't we ask the Superintendent to develop a comprehensive plan to improve student achievement and present the plan and the related costs next month?" -- are stated as questions but are also statements of belief of what should be done to improve the educational program and eliminate the condition that has created the perceived problem.

Chapter 3 takes each of these questions and explores why the question is misleading.

These questions or demands for action are based upon the erroneous assumption that a single factor is responsible for low achievement, poor discipline, low morale or whatever. The statement of the question commonly assumes that a quick and comprehensive change will be observed from the change of a single element, a single pronouncement by the administration, or the substitution of new materials or techniques.

This book suggests an alternative to the quick fix. We suggest that no single procedure, material, facility or person operates in isolation from the complex interactions among ongoing operations in the educational system.

In Chapter 4, six districts tell their stories. Each of these districts responded to the problem of increasing student achievement in a unique way. Each district, starting from a sense of district responsibility for student learning, evolved, over a period of tie, a method whereby the district used test scores and evaluative data not only to monitor but also to plan instruction; not only to diagnose but also to prescribe. Two of these districts developed centralized curricula-based systems where tests assisted teachers in increasing children's learning through a prescribed sequence of knowledge acquisition and skill development. Two others developed a feedback method, where central office staff interpreted test and evaluative data to school planning teams who advised on annual resource allocations.

Particular districts who approach the student achievement question in integrated organizational terms have developed such systems. But there are not many who do so. Why are these districts different from others? Or, to turn the question around, what are the impediments to all districts in doing the same thing? Why is instructional improvement through the management of testing and evaluation not a naturally occurring phenomenon.

Chapter 5, using the scenario and district stories, points out five reasons why a long-term coordinated district strategy is difficult to conceptualize. However, an open-systems perspective explains some of the barriers. It provides an understanding of how these changed assumptions can lead a school district in the direction of a coherent plan to interpret testing, evaluation and instruction.

Chapter 6 will personalize the foregoing discussions so that districts may begin to think about: 1) reasons for developing a unique system; 2) starting points, including analyzing opportunities and constraints posed by the environment and history; identifying existing ideas about testing, evaluation, and instruction; and looking at current curriculum, instruction and testing activities and what can be changed; 3) sequencing for development; 4) selecting emphasis (e.g., staff development, curriculum and materials, tests and evaluations); 5) identifying sources of support -- external, administrative and technical.

CHAPTER 2

School District Scenario

in which district personnel consider public reaction to test score decline as reported in the media . . .

CHAPTER 2

School District Scenario

It's Monday, well after 7:30 a.m. The district office of the ______ Unified School District is already full-staffed with personnel performing their designated functions. A new bus route just has been created for those students attending a self-contained class for the Educationally Handicapped. The Director of Food Services is dispatching the Monday menu of spaghetti, mixed vegetables, jello and milk to all of the district's kitchens. The clerk in charge of contacting teacher substitutes has made the usual wake-up calls and now operates the switchboard, third cup of coffee in hand. The Business Services branch hurriedly types a contractual agreement for today's staff development presentation.

The district is moderate in size with a diverse student population. Minority enrollment is on the increase, yet the total student enrollment is on the decline. There are fourteen elementary schools, two middle and two senior high schools. The staff has been stable over the years, but within the past two there have been many between-school teacher transfer requests. The teachers' organization endorses the transfers and is pushing hard to lower the administrator-teacher ratio and raise teacher salaries by 15 percent.

A few parents have already started calling the Superintendent's Office regarding last evening's newspaper article headlined -- "Secondary Scores Decline--Elementary Tests Remain Stable." Some of these callers are concerned about what the published percentiles mean for their children's chances to get into good colleges. Two citizens who had participated on the district's committee developing system-wide goals and instructional objectives, find it hard to understand that the district students haven't yet made any gains in reading and mathematics. One taxpayer suggests that dwindling federal dollars are being squandered through bureaucratic wrongdoings. One call is taken by the curriculum specialist who attempts to explain the scores in terms of changing student populations and the shifting district emphasis into bilingual and career education programs. A final phone conversation focuses on a related news article about the upcoming local tax election and the political involvements of some vocal Board member.

The Superintendent, Chris Dewey, faces another crowded day's agenda. The day-at-a-glance calendar sprawled on the desk indicates that the regular scheduled monthly meeting with his principals begins at nine o'clock. Pencilled in red for 8:45 this morning is a hastily called meeting with two school Board members who want some quick answers about the test score situation. At one o'clock, Chris will meet with the Executive Cabinet, which consists of the district's central office administrators.

On one elementary campus, some of the teachers convene in the lounge and discuss the newspaper story. A group of instructional aides hired through Title I funds are there too. Designated staff members are performing their normal supervision duties on the playground. The students seem happy. The majority of students are clustered in their own ethnic groups, moving about with lunches, instruments and backpacks. They play and chatter informally until the morning bell calls them inside to classrooms.

Reading lessons are conducted from 9:00-9:45 a.m. in all district elementary schools using the Sequential Stages Reading Series. At ten o'clock, the first of three recesses is held, followed by the Individualized Computational Mathematics Learning Program. Science, social studies and art are squeezed in as time permits and at the discretion of the teacher. The Curriculum Committee is concerned about this emphasis on reading and mathematics; the Board has established a policy that the most important basic subjects, reading and math, be taught in the morning when the students are fresh and alert.

Around the corner, the high school campus is quiet except for the few students attending a 7:15 a.m. Advanced Placement Class. Two custodians begin their work picking up litter after last Friday's football game with a cross-town rival. Six security guards, armed with walkie-talkies, have been called to the Principal's office. Due to power failure, no bells or clocks on campus are working. The guards will be used today to move students through the six-period day.

The high school's teachers' lounge is noisily filled with teachers criticizing not only the newspaper article but also the daily school bulletin, the teacher union president's position on hiring, and the rescheduling of today's sports rally. The reading and mathematics teachers are particularly concerned about the decline in the students' scores in their respective subjects. Just a year ago, the School Improvement Advisory Council set raising these scores as a top school priority.

There is much unrest on this campus. The high school, with its third principal in six years, has recently instituted its minimum competency exams with dismal results, and, there are rumors that the school may be closed next year due to enrollment decreases. The School Improvement Advisory Council has decided this year to tighten graduation requirements. Teachers and some parents are more concerned about what they consider the school's real problems, poor discipline and drugs. This year's projected staff development activities are intended to deal with several other issues: sex equity beyond sports and physical education, textbook and library censoring, mobilizing community resources, humanizing the schools, and teacher evaluations under the district contract.

The bull horn that has replaced the normal bells sounds the start of the first period and teachers slowly disperse to meet their Period I students. Ms. Swift, the newly appointed principal, experiences some frustration. Should she stay behind at school and deal with the bells or drive across town to the district office to attend the Superintendent's meeting? Does she have an option?

Parents and community persons are watching. This is her first principalship and Ms. Swift is the object of considerable district-wide attention. She is the first female high school principal appointed to the male-dominated district administrative roster. Her promotion from an assistant principalship in a nearby district has not been greeted with overwhelming enthusiasm by those who were comfortable with things "the way they were." She is aggressive, knowledgeable about new instructional developments and willing to try out some new ideas. Some members of the community and Board are delighted to see someone who wants to bring a different approach. Others feel that the district really needs funds to do the things that are tried and true. What is needed is more of the same.

Enough about Ms. Swift. She really had no option. She sped off in her car and was the fourth of 20 principals expected to meet at 9:00 a.m. with the Superintendent.

The school administrators sit in the staff lounge. Some are taking advantage of the coffee and donuts provided, some remark about yesterday's test results news.

It's 9:15 a.m. as the 20 principals remain in the lounge waiting unknowingly for a meeting whose entire agenda has been changed. The Superintendent is meeting with the two Board trustees. These two represent distinctly different community interests, but have a common concern about the district test results. The Board members insist that the Superintendent expand the Board agenda to include a statement from the Superintendent on the test scores, and a plan of how the district intends to raise these scores. Representing the voices of their constituents, they feel that the public is tired of excuses, and the Board members fear a swelling exodus of children away to the private schools. Likely the local television station will send a reporter to the meeting.

Mr. Pace, principal of the Bruner Elementary School for the last twelve years, arrives. He was stopped in his school corridors by Tim Huntley, a student teacher who wanted to know why Bruner had scored so low on the test. Mr. Pace tried to give a hurried but philosophical answer about the scores, making this comment: "However, young man, you must remember that our scores, at this school, weren't that low. In fact, we didn't change our scores. The papers are looking at our district performance on the average, not individually at our school. Look, I'd like to talk more, but I'm really late to an important meeting. Why don't you ask your master teacher?"

Huntley enters his 3rd-4th grade combination classroom rather perplexed. The students are working informally in semi-structured groups. They are beginning a humanities project he assigned which requires the exchange of ideas in order to develop critical thinking. The culminating activity is a group collage which captures the essence and tone of the topic, "Communicating Differences."

Not seeing his supervising teacher around, Huntly states that he has just seen the principal and that the principal permitted, in fact 'mandated' that they hold open discussion about the school, the programs and the tests that are given. Because the tests seem so meaningless to them, the students stop what they are doing and begin to chat and reminisce about who cheated, whose pencil broke, who marked the wrong grids, and who faked illness.

The supervising teacher, a ten-year veteran, hurries in with several copies of last night's neighborhood newspaper under one arm and dittoed reading materials under the other. She thinks she has a great idea for a new related assignment, "The Power of the Press in Communication," but instead, she verbalizes her pleasure with Huntley's initiative as well as the students' self-control in her absence from the classroom. The discussion over testing lasts about twenty minutes. Responding to the principal's suggestion, the student teacher asks his supervising teacher what she thinks about the test scores. The teacher looks puzzled, and is at a loss about how to reply. How much will her answer show that she really doesn't know much about the structure of the test and the meaning the different scores have for teachers? Hesitant and somewhat embarrassed, she begins a long discourse on the history of the testing program. "This district and community have engaged in an extensive testing program for the past fifteen years . . ."

Back at the district office the principals are admitted to the Superintendent's office at 9:30 and the meeting begins promptly. Superintendent Dewey deals with a somewhat routine administrative matter and reports on the just-concluded meeting with the Board members. "If I am to make a presentation to the Board on Monday I want to base my comments on the perception of people, like yourself, who have first-hand knowledge of what is going on. We've been pushing basics now for a long time -- yet the test scores remain at unacceptable levels -- what in the world is happening?"

A long, intense, but somewhat rambling discussion consumes the next two hours. The principals' perceptions of the problem cover many factors.

Some blame the influx of children with English as a second language and for whom the test is not a fair measure of their real skills -- or the schools' instructional program. Others think that the tests are meaningless to the kids and they don't try. Others insist that many students just don't know how to take such tests. Some respond with questions, such as: Does this test measure what we teach? Do we know which specific students or groups of students account for the lowest scores? Are all the teachers really implementing the new instructional program? Some suggest solutions such as more inservice training, or changing the tests, or not releasing the test scores.

The meeting ends with a marked feeling of general frustration. Ms. Swift, thinking about the forthcoming meeting, remarks that there are times when it is nice not to be a superintendent.

The Administrative Assistant to the Superintendent picks up the extra xeroxed copies, and gets coffee for her boss. It has been a busy day already.

It's 1:30. Chris is concluding a telephone call as the eight members of his executive cabinet enter the Superintendent's office. Sitting around the large executive table, from left to right, are: the Director of Curriculum, the Administrative Assistant to the Superintendent, the District Personnel Director, the Evaluation and Testing Coordinator, the Director of Pupil Personnel Services, the Director of Federal Projects and Programs, the District Budget Director, and the Associate Superintendent for Elementary and Secondary programs.

Staff assistants have brought additional school and district documents including last year's unused Title I Evaluation and Reporting System (TIERS), the basal readers used at each grade level, the mathematics learning packages, and the computer outputs presenting school and grade level information, as well as an item analysis for each question administered to the students.

The Superintendent has described the unexpected meeting with the Board members and the meeting with the principals. The chief administrator outlines what the Board is expecting and he asks them one by one to give his/her perceptions on the problem and any ideas he/she might have both as to solutions and as to how to handle the Board presentation. But not everyone sees the problem the same way.

The Curriculum Director starts off. "I hope we don't panic over these latest test scores. After all, our new curriculum emphasizing basic skills has only been in operation for a little more than one year. You can't expect that things will change overnight. These things take time. My advice is to keep emphasizing what we are doing -- we have a sound plan for getting kids to study more basic skills. My advice is to say that this is a temporary condition -- things will improve with time."

"Besides," the Curriculum Director continues, "I'm being plagued by self-appointed book censorship groups who are insisting that we clean out our texts and school library of what they consider offensive materials. When this issue reaches the Board -- then we will have a real problem!"

The District Budget Director, an accountant by training, sees the problem as one of ineffective teachers who continue to do an inadequate job in the classroom. "Why not," he argues, "begin aggregating these test scores by teacher and see if any teacher continually has classes of students who fall below a pre-selected percentile? These teachers, once identified, could be given extra help and if they still perform below district standards, then we can 'counsel them out of teaching.' In the business world we do this all the time -- I mean, we look at a salesman's monthly records. If he can't perform -- out he goes! All the new curricula and regulations and media don't add up to anything if you have weak teachers. I say, use the data to attack what is obviously the most serious problem -- some poor teachers!" Most in the room shift uneasily in their chairs.

The Coordinator of Testing and Evaluation attempts to cool what might become a heated discussion. "At the risk of repeating myself endlessly, let me say that I warned you some time ago that this might happen. As you will recall, I predicted that there was a mis-match between our present norm-referenced tests and our newly developed curriculum. As it is now, the tests simply do not measure what we are teaching. At the time we changed our curriculum we should have looked at a different, more relevant norm-referenced test. But, I'm afraid it is too late to do much about it right now. Any shift we make will immediately raise suspicions that we are selecting an 'easier' test, just to make us look good."

"Having said that, let me suggest a more immediate solution, namely teaching the kids to take tests better. I don't mean anything illegal, like giving them the test answers. Don't get me wrong. But many teachers report that some kids are really frustrated by the whole process. Indeed, some kids are forced to take the test even when they can't read English. I ask you -- of what use are those results? What are we doing! Also, some kids are confused by the standard answer forms. I say let's have some teacher inservice training on how to teach kids how to take tests, and let's remove those test scores of kids who obviously cannot read English very well -- or better yet, let's get the test translated into their language. These kinds of actions will bring more immediate improvement -- something we desperately need."

The Superintendent's Administrative Assistant, who normally doesn't say too much at these cabinet meetings, unexpectedly speaks up. "It seems to me that we are not the only district with this problem. Why don't we find some other districts that are like ours in size and student make-up, etc., and see what they have done? If they have come up with something that works for them -- let's adopt the same plan here. Such an approach will save money, time and will be successful. Those are words that the Board likes to hear."

Finally, the Associate Superintendent for Elementary and Secondary Schools speaks up. "At the risk of further complicating what has been said and suggested, let me remind you -- us -- that the problem is not equally distributed in the district. Our elementary school scores did not decline but the secondary ones did. Also, some elementary schools actually increased their test scores. As we struggle with this problem we should be mindful of these facts. No simple, single strategy is going to solve this problem, even if the board and some segments of the public are clamoring for one. We must very carefully think through a long-term, comprehensive strategy that will get at the core of the problem -- otherwise, we will be holding this same discussion again this year -- that is, if we are still here."

The Budget Director reminds the group that any plan must be developed within the confines of yet another budget cut. There will be even less money for staff development, evaluation studies, and the like.

At this point, the Superintendent turns the discussion to the district enrollment and budget projections and the administrators submit plans they have developed to accommodate the anticipated cuts.

At 4:30, the Superintendent thanks them for their perceptions, ideas and advice, and ushers them out of the office. The secretary buzzes and reminds the Superintendent that he will have to leave in an hour in order to make a presentation at the district's student citizenship dinner.

It's now 9:15 p.m. At home, the superintendent sits down, legal pad in hand, to sort out what has been said. The several people to corner him at the citizenship dinner about the test scores made it clear that a potential crisis is brewing. What he says at the rapidly approaching Board meeting takes on increasing importance. He looks at his notes, searches his memory and begins to categorize the suggestions made. The following appear.

1. The test approach
   - select a different test
   - teach the kids to take tests
   - relate the test content to the curriculum
2. The teacher approach
   - improve teacher skills
   - identify incompetent teachers -- and provide inservice training
   - fire those who can't meet district standards
3. The planning approach
   - develop a comprehensive plan that considers the several dimensions of the plan
4. The adoption approach
   - identifying districts that are similar to ours
   - sort through their respective plans and pick one that has been effective elsewhere
   - implement it in our district

Somewhere among all these approaches, there lies an answer to the district's dilemma -- but what is it, and how does he begin to find it?

Discussion

While the details might differ, we suspect that there are characteristics or parts of this scenario that are familiar to many administrators in school districts across the United States --

- districts that collect test data and find those data released to the public through the media;
- districts in which there is a public outcry when the test data do not reveal adequate pupil learning and achievement;
- districts that are urged, or motivated to "do something" to bring up the test scores;
- districts where there are multiple and differing perceptions about what causes the problem and, accordingly, about what will best achieve the desired goal of raising test scores.

Such districts face a very real challenge and dilemma because they are often confronted with data that challenges their credibility to perform their central instructional role, because it is widely assumed that test scores accurately reflect the school's instructional program. The public often assumes that there is some causal relationship between the district's instructional program and test score fluctuations. Often the public demands, and the district officials promise, that the district will make necessary instructional program adjustments so as to reflect increasing rather than declining pupil achievement.

But what should be done? This, of course, rests on knowledge about the district's instructional program and its relationship to the tests used to measure instructional effectiveness. In many districts this relationship is simply unknown or only vaguely understood. The districts have not examined the relevance of the testing program to the instructional program. Indeed, many districts would be hard pressed to describe the degree to which their instructional program is being taught by district teachers behind their classroom doors. Thus, the districts, such as the one portrayed in our scenario, can be characterized as possessing instructionally relevant data but data that, in its present form, is not being adequately utilized as a potential management tool.

This book is intended to help school administrators, board members and teachers to better understand the potential that district testing and evaluation programs have for improving the instructional program. Actually, our purpose goes considerably beyond just focusing on testing and evaluation as isolated school district functions. We conceive of testing and evaluation as being components of a district instructional management information system.

To elaborate, we define a district instructional management information system as follows:

District - We consider instructional improvement as a school district responsibility. To be sure, the individual teacher in the classroom is the basic component of the instructional program, but this does not relieve the district from full responsibility of seeing to it that the teachers are competent, using appropriate instructional methods, and furnished with necessary supplies, equipment, and instructional materials. Teachers alone cannot carry the entire instructional burden; neither should they assume the total responsibility and blame for the instructional program's shortcomings. This responsibility rests with the school district.

Instructional - By this we mean the delivery of the district's curricula. It includes the teachers' methods of determining learner needs, instructional decisions and activities, and assessments of the effectiveness of instruction, including district-wide testing and instructional program evaluations.

Management - This includes the planning, developing, implementing and assessing the district's instructional programs. It involves coordinating many components ranging from personnel to budget to school site administration to inservice training.

Information - This includes data collected about such topics as pupil achievement, teacher behavior and activities, program implementation. It includes using many kinds of tests, program evaluations, opinions and attitude surveys.

System - This recognizes that the various components, e.g., instruction, test scores, evaluations, inservice training, have a relationship to one another and that they should not be isolated one from the other. For example, a change in the instructional program in a given subject will, or should, have some impact on other components, e.g., tests, inservice training.

Put another way, we propose that districts can more efficiently and effectively manage their instructional programs if they have developed a system for relating information and data they have gathered to their instructional program.

Our scenario district is rich in data, but poor in not having developed a management system that will allow them to use those data for instructional program improvement. This book is about how such a district instructional management information system can be developed.

We think such a book is needed at this time. Schools and school districts are not doing very well. That is, public schools are increasingly viewed by the public as not being very effective instutitions. While much of the present crisis in public education results from nationwide, indeed, worldwide, economic decline and political priorities that have shifted away from public education, one can argue that public schools must present evidence that they are effective organizations before the public will express much confidence in the schools. A popular public perception is, that merely "throwing money" at the schools will not necessarily improve the quality of the schools.

Experience has shown that single element strategies just don't work. For example, some have advocated massive inservice teacher training. This strategy assumes that the individual teacher behind the classroom door is the problem. There is some merit to this approach, but in our view, it is too narrow. It does not address the systematic and district-wide responsibility for instructional improvement.

Others have advocated a data-based accountability approach in which test data are used as a prod to get teachers to be more responsible -- and therefore, ultimately more effective. This hasn't worked either; it has resulted in the mis-use of test data as a punitive device which has resulted in deception and evasion and a lessening of trust between teachers and their supervisors. Again, this strategy is not based on any comprehensive, systematic instructional improvement system.

We believe that what is needed is a system or approach that will link together our advancing knowledge about testing, evaluation and instruction within the complex setting of school districts. If done properly, it can provide one approach to improving school districts' instructional programs. Clearly, it is not a panacea. It has its limits. For example, there are important school activities and outcomes that are very difficult, if not impossible to capture or measure in a testing program. We are talking about such outcomes as the students' self-confidence, or appreciation of beauty, or enthusiasm for inquiry. Thus, one has to be careful not to begin diminishing the provision of these important activities on the basis that they cannot be adequately measured, at least not with our present level of testing technology. Frankly, we doubt that some of the schools' most important outcomes can ever be measured, except in a very subjective way. But this doesn't lead us to conclude that all measurement and evaluation of performance is undesirable. We must do what we can, be aware of the limitations, and take actions necessary to see to it that the schools more clearly realize their instructional potential.

CHAPTER 3

School District Questions

in which four simple questions implying "quick fix" solutions are reframed for more productive thinking . . .

CHAPTER 3

Introduction

The reader will recall that we left the superintendent of our scenario school district at home, late in the evening, legal-sized pad in hand on which he had written the four approaches he had heard about the looming crisis during his conversations with his principals and his cabinet. The four approaches were: the test approach, teacher approach, planning approach and adoption approach. Our superintendent, and others who may find themselves in a similar situation are indeed faced with a formidable challenge. The crisis is real. That is, the public believes that, at least as measured by the district's testing program, the district simply isn't educating children to acceptable levels of learning. What is more, the decline in test scores seems not to have been stopped by the efforts already undertaken by the superintendent and his administrative and teaching staffs who care about the same problem. Not only do they care, they have acted. They have formed a district community/parent committee to define district goals -- perhaps in the hope that if the goals are more clearly explicated, they will also be more obtainable. Yet, in spite of concerns and effort, the district remains plagued by unacceptably low test scores which presumably represent weak student achievement.

But the superintendent's dilemma is not only that the district's efforts in the past have not worked. He is also challenged by the four suggested approaches he has recorded on his note pad. Each has the "ring of truth: about it. Does the answer to his district's dilemma rest on following one of these approaches? Or some combination of approaches? Or in going in an as yet unidentified direction? The board expects an answer. He and his staff want an answer. Where does he begin?

We suggest that he begin by understanding the myths and the facts underlying each of the questions.

J. Richard Harsh, the author of this chapter, draws upon considerable training and experience in tackling the difficult task of analyzing these approaches fairly and concisely. For many years, he was the California field director for Educational Testing Service. This brought him into a working relationship with many school district instructional, testing, and evaluation programs. Since retirement, he has served as a much-sought-after consultant who specializes in district testing, evaluation and instructional problems and programs. In writing this chapter, he has drawn extensively upon his many years of experience and the insights they have provided.

I. How Soon Can We Expect Test Scores To Go Up?

The publication of district test results in the local newspaper often sets off a strong reaction in the community. The board of education may receive strong demands from vocal citizens to initiate reforms in the instructional program that would improve the achievement of the students. Special funds may be allocated to schools to improve the test results. Such a flurry of activities stimulates people to question past and present practice and ask what results may be expected.

"How soon can we expect the test scores to go up?"

This question can't be answered without knowing which tests are referred to. However, as this is mentioned, the quick retort is often -- "the tests in reading, mathematics, and language that have always been given."

Myth: All tests with the title of "reading test," measure reading in the same way, or, all the tests of mathematics measure the same skills and knowledge of mathematics.

Fact: There are a variety of tests that are given to measure students' skills in reading, math and language, and the various tests (by design) have different relationships to the local instructional program. Various tests of reading, math or language may ask students to demonstrate different skills, knowledges and applications.

Discussion: There are many categories of tests. Each category of tests reveals different things about student learning. And, within each category of tests, the tests may differ from one another. The following array of reading tests may be used in a school district.

Published Standardized Reading Tests

These tests are developed, normed and published by commercial test publishers for sale to educational institutions throughout the nation. They provide norms constructed by the publisher that are derived from testing a sample of students in various types of school systems that the publisher believes to be a representative population of the students in the nation. The questions on these tests are designed to measure the skills, knowledges and applications that are believed to be most generally common to many reading programs throughout the nation. These tests may be considered as measures of generally common reading skills with norms that suggest the ranking of levels of performance on the test.

The California Achievement Tests, the Stanford Achievement Tests, the Iowa Tests of Basic Skills, the Sequential Tests of Educational Progress, the Metropolitan Achievement Tests, are examples of these standardized tests that have sub-tests to measure the skills of reading, mathematics and language that are considered to be generally common to many school programs. Analysis of the reading tests of these publishers reveals that the content and skills measured by the questions differ from one test to another. For example, reading vocabulary is measured by some tests (CAT, SAT) by presenting a word and asking the student to identify from among four options that one which means the same as the word given. In contrast, the Iowa Test of Basic Skills measures reading vocabulary by asking the student to identify the meaning of a word in a particular sentence or story. E.g.,

> **Bob said, "I'm glad that I can bank on your loyalty."** In this sentence the word BANK means: a) a place to keep money; b) the side of a stream or river; c) to be able to depend upon another; d) to arrange a stove or fire.

State Assessment Programs

These tests are designed under the direction of state educational agencies to provide measures of the skills, knowledges, and application that the state requires in the curriculum or that are expected to result from the instruction in all districts of the state. Some states may construct their own tests (California, New York, etc.) while other states may contract with commercial test publishers to develop tests to their specifications (e.g., Florida, Michigan, etc.).

Norms for these tests are developed from the performance of the students in the particular state. The norms may present numerical ranking of schools within the state, or percentile ranks for different levels of performance by individual students, or may provide numbers and percents of students within each school or district that demonstrate competence in each skill according to the standards established by the state agency. Some states (California) provide norms for groups of schools with similar demographic characteristics such as mobility of students, socioeconomic conditions, percent of minority population, etc. The different norms both within and between state assessment programs provide different information regarding the performance of individual students, sub-groups of students or the schools and districts being assessed.

The use of the state testing results also differs in relation to the design of the tests and the manner in which students are tested. In some states, all students in all schools take the same test on the same day of the year. In contrast, the California State Assessment program is not designed to provide information about individual students. California uses a "matrix sampling" design which provides many different short forms of the assessment of basic skills in reading. The combination of all forms of the test covers all the skills that are desired to be measured, but any single form will measure only a portion of the skills. Students take different forms of the test that are randomly assigned within each school. The scores from all forms of the test are reported for schools and districts but no scores are available for individual students. The report made to districts in California provides a "percent of correct answers" to the composite of all skills measured by all forms of the tests. Such a report provides information of the percent of students in a school or district that are successful with each skill measured. This information is quite different from the state reports that provide percentile norms for each student or school population, and neither of the state reports may be compared to the national publishers' norms.

Proficiency Tests For High School Graduation

By 1980, thirty-eight states had adopted legal requirements of proficiency testing prior to granting a high school diploma. The designation of competencies and the standards of performance were made by state departments of education in some states while other states mandated local definition of the competencies to be measured and the level of performance required for high school graduation. The most common practice was to consider these assessments as measures of the minimum essential skills in the language arts and mathematics that would be required by out-of-school daily living. In California, where the law required local definition of competencies, such designations were made by district committees of citizens and educators.

The tests commonly measure skills in reading, math and writing that are conceived as essential for coping with post-secondary life. Thus, the content measured by these tests typically covers only a small portion of the reading, math and language curricula of a district.

A wide variety of instruments have been used to assess the minimum essential skills for high school graduation. When a common test is required by the state department of education, some states constructed new instruments, contracted with testing companies for unique development, or adopted a commercially prepared test of basic literacy. In those states such as California where the responsibility for the assessment was delegated to the local district, a wide variety of tests have been constructed locally, or groups of districts have agreed on common definitions and constructed, contracted for, or purchased commercially prepared tests.

The proficiency tests for high school graduation vary widely between and within some states both in terms of the skills measured as well as the level of performance that is required. In the context of local definition, the test results are meaningful only in terms of the particular skills and levels of performance that have been adopted. The results of these tests thus may have very different relations to the local curriculum, the state curriculum, or nationally standardized tests of comprehensive measures of the common instructional program.

Due to the fact that the legislation requiring high school competency examinations frequently assigns penalty to students not performing at a specified level (denial of diploma), high attention has frequently been given to devising instructional packages to insure that the particular skills measured by the test were mastered by all students. Increases in student test scores in this context may be achieved more rapidly than increases in test scores that assess the full spectrum of the school curriculum.

Elementary Proficiency Tests

These tests have emerged in many states and regions as extensions of the legal mandates for demonstrated competence in the basic literacy skills. As secondary districts were confronted with the wide range of academic achievement of entering students there was concern expressed for the academic development attained in the elementary schools.

A range of strategies for the definition of essential competencies includes local district, counties, states, and regions specifying contents and levels of performance that should be attained by the middle or upper elementary grades. In some regions of states such as California that require local definition, there are attempts to articulate required minimums from the elementary through the secondary schools. The concept behind the design of these tests is that certain essential skills should be attained by all regularly progressing students by a particular age or grade.

Generally, these tests, as is true of the high school graduation tests, are designed to measure the essential skills in reading, math, and language that are locally defined as necessary for continuing educational progress. The questions are of a limited range of difficulty that assesses a particular level of competence expected at the end of a particular educational period, such as sixth grade. The tests do not measure the full scope of the curriculum in mathematics or language arts, but provide monitoring information concerning the number of students that have achieved the essential skills by the upper elementary grades. The fact that districts may define the skills and criteria of mastery makes it difficult to compare the performance on these tests with nationally normed tests, or state testing results.

Curriculum-Referenced Tests

These tests are intended to measure student mastery of the content and skills of a particular curriculum. Some of these tests are constructed by the publishers of curriculum materials to measure the mastery of that particular content. Other tests may be constructed to measure the curriculum developed by a local educational agency. The locally constructed measures are designed to assess the student attainment of the objectives of the local instructional program. A wide variety of criteria are used to interpret the scores made on the curriculum referenced tests. Some use the percent of students answering each question or group of questions measuring a particular skill while others develop local percentiles to array or group the performance of the student populations.

These tests are probably most directly related to what students have experienced in the instructional program. Insofar as the tests systematically assess the content and objectives of the local program the results may be considered to be relevant to the question of the degree of mastery that students have attained. On the other hand, the local program may have unique contents, instructional materials or unique sequence and emphasis on certain skills that may diverge from what may be emphasized in another district or what may be measured by external measures such as state or nationally standardized tests.

Tests Differ From One Another in What They Tell About Student Achievement

- Nationally standardized tests are assessments of broad general outcomes that are conceived to be quite common to many educational programs. The norms provided by the test publisher provide a method of ranking local student population's performance in relation to a sample of students thought to be representative of the national student population.
- State testing programs are designed to measure those skills and knowledges that are judged to be the common desired outcomes of instructional programs in the state. The norms or summary statistics provide a means of comparing local student performance with the total student population of the state.
- Proficiency tests, especially those required for high school graduation, tend to emphasize minimum essential skills that are thought to be necessary for the demand of post-secondary school life. Student performance is interpreted by applying local standards of the district or the standards established by state mandates. Elementary proficiency tests tend to reflect local designated skills and knowledge that are deemed essential to further educational progress.
- Curriculum-referenced tests measure the attainment of the content of a specific curriculum or local instructional program. Interpretation of the student's performance is made in relation to the extent to which the student has mastered those objectives of the local instructional program.

"These explanations are informational but the question originally asked still persists -- What can we do to make the test scores go up? The low scores that have been reported for the district tell us that the students aren't learning what they should!"

Myth: If the group scores of a district go down, it automatically means that students are learning less than in previous periods.

Fact: Many different factors can account for changes -- either up or down -- in test scores of the student population of a district.

Discussion: District test results are usually reported as the average score of all students such as the "38th percentile," or "forty percent of the questions were answered correctly," or "the district ranks below sixty percent of the districts in the state." Such summary characterization of the performance of the total student population often conceals more than it reveals. These quantitative summaries of the central tendency or average of the group do not tell the number or proportions of students making scores well above the average reported, nor do they tell about extremely low scores that may have great effect on the average of the total population.

Persons with anxious or critical disposition are prone to take the "average score" as the characterization of all students. It is possible that the average score reported was made by no student. (I have two sons. One is 5 feet tall, the other is 7 feet tall. The average height of my two sons is 6 feet -- this doesn't provide an accurate description -- only a numerical characterization of the average height.)

However, regardless of the method of reporting test performance, test scores do go up and down. With the changes in the reported scores it is appropriate to ask -- "What are the factors that might change test scores."

The mobility of the population within regions and the nation produces rather dramatic changes in the characteristics of the students in many districts. As occupational opportunities shift in location there are related shifts in the population which works in the occupations. The development of a large technical industry in a particular community may bring a new population of professionals and technicians. A residential community that is changing to diverse industrial plants may have a shift in population as the workers required may have different education and socioeconomic status from the former population.

The changes in the population of a school attendance zone may produce shifts in the abilities and achievement of the student population. For example, in a period of two years, a school district increased its enrollment by 25 percent and also observed an increase in the median prcentile score in reading from 51 percent to 64 percent. During the same two years, another district in the region experienced 50 percent of the population moving to other locations and a new population of equal size but different demographic characteristics moving into the district. In this district with high mobility the test scores showed a substantial decline.

Test scores may rise or decline because of a change in the tests that are used to measure the basic skills. The change in test scores may be related to either variations in the norms of the tests, variations in the types of questions used by different tests, variation in the content and skills required by the test questions, or a combination of all of these factors. Such changes in test scores are often related to the degree to which the skills measured by the test are congruent with the content, sequence and emphasis of the local curriculum.

Pressure to raise test scores has sometimes resulted in inappropriate activities. As mentioned previously, the summary report of test scores typically presents the average of the scores of the total student population. Test score averages may be raised by the unprofessional action of "selecting a sample" that omits those students who are known to be low achievers. A similar distortion of the average test results may be produced by the lack of effort to insure that all students are tested. A regretful result of such a strategy revealed that only two thirds of the population of a high school were present for the test administration.

The concern for raising test scores has produced a wide variety of procedures to prepare students for the test they will take. The most inappropriate procedure is teaching the exact content and format of the questions in a standardized test. When persons have been accused of "teaching to the test," some have defended their actions by saying that they were merely using practice exercises to develop student "test wiseness" or familiarity with the format and the type of questions contained in the standardized test. While there may be appropriate activities to familiarize students with test-taking skills, there is no question of the inappropriateness of revealing and teaching the actual content of test questions. Such actions reflect a total lack of understanding that the questions on a test are only a sample of a very large domain of skills, content and application. The teaching of the particular content of test items that are to represent an entire domain leads to the consequence that that the test scores are erroneously raised. The students are further injured by the assumption that their test scores suggest that they have achieved particular skills and knowledge of a domain when they have not.

The consequences of strategies designed to rapidly raise test scores by inappropriate and unprofessional activities are very serious for both educators and students. Moreover, the accumulated data base on test performance that currently exists suggests that the inappropriate activities will be discovered within a period of weeks or certainly by the next year.

The combination of sincere and appropriate efforts to improve student achievement has high likelihood of bringing about a rise in some test scores. At the same time, it is useful to consider whether the rise or decline of the average test score has any predictable significance for individuals in a particular class in one school, for boys or girls, for students with limited English proficiency, or for particular skills such as vocabulary, solving problems, or applying the skill in out-of-school situations.

In a school or district with several hundreds or thousands of students, an increase in the average score provides little assurance that a particular student's score will be higher or lower. For example, concentrated effort with the most severely retarded students may result in an increase for them, while the most able or accelerated students may show no gain or even a decline. When an average is computed for a large population the increase of tests scores for those who had previously made very low scores will raise the average score but there may be no improvement for the more able students. The converse may also be found: emphasis given to the group of high achievers may raise their scores while the less able show no improvement or a decline.

Sub-groups within a school or district student population may not demonstrate the same increase or decline that is suggested by shifts in the average score. It is not uncommon for individual schools in a district to have average scores that are above or below the average score for the district, and ethnic sub-populations often show substantial divergence from the district average.

The increases in average scores in reading, math or language do not provide assurance that all skills within these areas of the curriculum have improved. It is quite common to observe an increase in the average math scores following concentrated intervention or emphasis on math computation. However, it is often observed that while the computational skills have improved, there may be no comparable improvement in the ability to solve word problems or apply the skills in out-of-school situations.

These examples of the many factors related to the rise or decline of average test scores suggest that careful analyses of the congruence of the content measured by the test and the local curriculum as well as systematic analyses of the needs and characteristics of sub-groups of the total student population are essential for understanding the meaning of test performance.

Such explanations are helpful to incease understanding, but the interested parties involved in education will persist with the conviction that improvement in the educational process will improve children's learning and these changes will be mirrored in the improved test scores.

Myth: We can dramatically affect children's learning and have the change reflected in rising test scores if the schools do things right.

Fact: This is a true myth. However, it is a very complex process to enhance student learning in the comprehensive goals of the curriculum and demonstrate that change through test scores.

Discussion: Longitudinal studies of the acquisition of skills and knowledge of the population served by the schools suggest that very great improvements have been achieved since the beginning of the twentieth century. Statistics on literacy, the knowledge explosion in the sciences, the field of automation and the numbers of students graduating from the various levels of education attest to these improvements.

There is need to recognize that education is a process that is constantly being affected by social and cultural forces that restrict or facilitate the learning of individuals. The learning of an individual includes the internalization of interests, work habits, aversions, attractions, values, social participations, as well as the skills, knowledges and applications of the academic curriculum. Individual learning is unique to that person and each person has pressures and demands that establish priorities and time schedules for the various skills, knowledge, and participation that will be attended to within any time period. The construction of an instructional program that will be effective for the diversity of a heterogeneous student population must be planned to enhance individual learning with all the diversities that naturally exist between persons.

Human learning cannot be designed as a factory might design the processing of inanimate materials into products such as cars, can openers, or washing machines. Thus, while there may be often-stated broad goals that are to be realized through education, the translation of these goals into the relevant contents of study and the design of instruction for individuals presenting wide differences at entry to the school classroom creates a very complex task.

The assumption that what a student knows and can do may be directly assessed by a test of a sample of questions of a very large domain is open to many reservations. All persons can recall circumstances in which they failed to answer a particular oral or written inquiry within a field of knowledge or endeavor that they were functioning in with satisfaction and commendable productivity. Such occurrances are similar to the concerns described in a previous section with the congruence of the content and format of test questions with the curriculum and instruction the student has experienced.

Enthusiastic support may be given for the assertion that children's learning may be improved through careful planning and implementation of an instructional program. The assessment of the outcomes anticipated from the program must be designed to insure that there is reliable and valid evidence of the attainments of the content and outcomes of the program the students have experienced.

Identification of the strengths and weaknesses of the various contents of the curriculum, of various sub-groups that compose the total student population, as well as existing and desired strategies of instruction are important actions to precede the development of an instructional plan that will improve the tested performance of the students in the district.

II. Why Don't We Identify Incompetent Teachers and Do Something About Them?

This question is frequently asked of administrators by board members, district advisory committees or individual parents who are concerned with a wide range of conditions that might range from increasing school budgets, vandalism in the school and community, low test scores published in the newspaper, discipline problems in the school, reports of academic deficiencies of local students applying for college admission, etc., etc.

The statement of the question frequently does not identify the problem or condition that will be solved by "doing something about incompetent teachers." Moreover, there are varieties of sub-elements or sub-questions that need clarification prior to a response or development of a plan for discussion or action. Certainly an initial step for clarification is the recognition that the question as stated asks for two actions -- identifying incompetence and then doing something to eradicate it. Identification of competence (or incompetence) of teachers requires a detailed analysis of the personal and professional knowledge, skills, attitudes and behavior that are needed to implement the instructional program specified by the local board of education and administration. It is also apparent that this request to eliminate incompetent teachers stems from the belief, by the person asking the question, that an undesirable event or condition is directly and solely the result of certain persons' incompetence. Of equal importance in responding to this question is the need to identify the evidence or criteria that the persons asking the question are using to draw the conclusion that there is incompetence among the teachers.

The statement/question -- "Why don't we identify incompetent teachers and do something about them?" -- needs additional discussion and clarification to identify the problem and the evidence that prompted the question. Since there is no generic list of characteristics that identify an incompetent teacher, answers to further inquiry are indicated. To examine the postulation that there are incompetent teachers the following questions may be in order.

- What are the indications of an incompetent teacher?
- Is the incompetence only noted with students of particular achievement, motivation, cultural background or age?
- Does the lack of competence appear to be found at certain grade levels or with certain subjects of the curriculum?
- What criteria or evidence is available and used to document the competent and the incompetent teachers?
- What causes children not to learn what is expected?
- Are there legal, political or institutional constraints on the actions that may be taken to deal with the incompetent?

The response to inquiries for clarification of the problem and the evidence or condition that promoted the question of dealing with incompetent teachers often includes reference to the reported test results of the district. This was the case when a district advisory committee was formed to investigate the continuing decline of the test scores of the local schools and presented the following questions to the superintendent of schools.

> "Can you obtain the test results for each teacher's class or classes?"
>
> "Can you compare the results for all teachers and then determine which teachers get *high or low achievement?*"
>
> "Wouldn't this be a good way to identify the incompetent teachers?"

These questions concerning the use of test results should prompt the further question: Can -- or should -- test results be used as a single criterion of effective teaching? Knowledge of the multiple causes for behavior and learning, coupled with the longitudinal accruals throughout a person's educational experiences indicates that although test results might be used in this manner it would be very inadvisable. However, persons proposing such a criterion for identification of competent teachers do not readily accept the reply that such a procedure would be inappropriate without further explanation. Examination of the proposal indicated by the question reveals several myths that provide the basis for the erroneous assumptions.

Myth: This year's test results for each class or grade are a direct indication of what the teacher has taught and what the students have learned during the year.

Fact: There are many different types of tests which may be inaccurately related to what has been taught in a particular class or school.

Fact: Most tests measure what has been learned in previous years as well as those outcomes that are expected for a particular age or grade.

Myth: Standardized norm referenced tests indicate how well the teacher has taught and how well the students have learned common skills and knowledges.

Fact: Nationally standardized tests have no exact match to what each teacher in each school has taught.

Myth: The teacher having a class with the highest test scores is the most competent teacher.

Fact: Student populations may have very different levels of achievement at entry to a particular teacher's class. The test results for one particular year are not solely a reflection of what has been achieved in that year.

Discussion: Interest in using test results to evaluate the effectiveness or competence of teachers is plagued by the inexact match between what the teacher has taught and what is measured by the tests. It is well known that the nationally standardized tests are designed to measure the outcomes of instruction that are generally common to most districts throughout the nation. Thus, while there is a general match for the most common outcomes of instruction there is an irregular and poor match with the content of instruction in some classes and schools of the nation. Moreover, while the general outcomes may be common goals for many or most schools in the nation, there may be wide difference in the sequence and timing of the particular skills that are required to attain those goals.

The use of locally designed tests that are intended to measure the objectives of the local curriculum does not offer assurance that the results of any single testing will identify effective and ineffective teaching. In any school or district rather substantial differences may be found between classes in the student entry characteristics to that class. If a particular teacher has a substantial number of students who have retarded language development, come from homes that provide marginal support or encouragement for academic achievement and

> suffer from a cultural disadvantagement it is very unfair to compare the test results of this class with another class of students that demonstrate normal language development and strong support and motivation for academic achievement.
>
> It is recognized that tests properly selected to assess the desired outcomes of instruction of the local curriculum are very useful for instruction. Such diagnostic use of test results often provides specific direction for a teacher to emphasize particular contents and skills with that class, while another teacher may need to provide emphasis with other skills and knowledge are most appropriate with that class of children. The use of the same test results with both classes at the end of the year can not be just globally compared in terms of which class had the highest mean score, but should be analyzed in relation to the progress the students have made during the targeted instruction of the year.

"Alright, it has been recommended that test results should be used to diagnose the strengths and weaknesses of what children have learned. Why not list the weaknesses that show up in each teacher's class and then give the teachers some training in teaching those things the children haven't learned? Parents and students agree that there are good and poor teachers in every school. This is just a request to train the 'poor' teachers to become 'good' teachers, and even give them some incentives such as a bonus or a raise. If the teachers don't improve with the inservice training then the district should fire them or transfer them to other work."

Several myths are behind these comments.

Myth: Student learning is solely related to whether they have been taught particular content during a particular year.

Fact: What and how students learn is the result of what they bring with them to the classroom, what their peers are doing, what the home does to encourage and support learning, as well as what is organized and presented in the instruction.

Myth: "Poor" teachers lack the knowledge of the subject matter and can become "good" teachers by having more training with the content of the subject they are teaching.

Fact: Instruction has many facets that relate to the diverse characteristics of a heterogeneous group of students, the attitudes and work habits that may or may not be developed by the students in previous education.

Myth: There is a single best way for all teachers to teach all students.

Fact: Children have different styles for effective learning. Teachers also have different instructional skills and different styles for effective communication and motivation of students. The same method of teaching and the same content will not be equally effective for all students in a class.

Myth: Labeling "good" and "poor" teachers is an effective technique to improve instruction and staff morale.

Fact: Focus on personal inadequacies tends to produce negative or defensive reactions. Public comparisons of staff produces alienation and distrust among staff.

Myth: An ineffective teacher can be "fired" or transferred at any time by the administration.

Fact: There are powerful legal and political constraints on the firing or transferring of professional school personnel. Teacher contracts and contracts with the teachers' professional organizations require lengthy due process with detailed documentation that must be proven to be reliable and valid evidence for action.

The suggestions of demands for the identification of incompetent teachers and the institution of remedial action or dismissal are generally related to the assumption that the learning of students is directly related to the competence or behavior of a particular individual. This assumption ignores the reality of the school functioning as a dynamic system that is influenced by the combined contributions of all the staff, administration, board of education and the community.

The purpose of citing some of the myths that surround the demands for immediate action to improve student learning is not to deny the constant need for improvement or to offer rationalizations as to why change is impossible. In contrast, these illustrations are offered as evidence of the complexity of human learning and the complex relationships that exist between the school, the staff, the children and the parents of the community.

The identification of the personal and professional characteristics of effective teachers is a worthy yet very complex undertaking as has been summarized by decades or research by such persons as Michaelis, Ryan and Gage. Thse extensive studies of teachers and teaching provide evidence for the unproductive results that may accrue from efforts to change individual persons or single elements of the educational process. Rather than giving support to the quick and parsimonious demand, it is more profitable to consider the answers to questions concerning the multiple causes of behavior (and learning) and the operation of the school within the content of the community it serves.

The request to identify incompetent teachers and then do something about them must also be considered in relation to the conviction that the teacher is the key element in the facilitation of student learning. The foregoing discussion of the myths and erroneous assumption regarding the use of a single evidence of incompetence, such as average class test scores, suggests that a comprehensive evaluation of the processes and product of the instructional program be made. The evaluation of the teacher must be made within the content and expectancies of the teacher in the organizational system of the district. The case studies presented in Chapter 4 suggested that the teacher's role (and direction or supervision) is different in the highly centralized system of goals, objectives, materials and methods being decided and managed by the central office as contrasted to the system that relies on decision regarding conduct and sequence of instruction to be made by each teacher in the classroom. A different role and process of instruction may be found in the system that charges the school principal with the responsibility of designing instruction that will be most effective for the students of that particular school. The three organizational systems found in the case studies suggest that the role and expectations of teachers may be quite different in the various systems.

Recognizing the varying role and expectation of the teacher in different schools and systems suggests that the criteria of teacher competence must relate to the local system of governance and supervision, and the expectancies that are held for teachers within the local system and conditions. In effect, it is essential that the school system (and the teachers) reach explicit agreement on the role, responsibilities and outcomes that are anticipated.

A different perception of the role, responsibility, behavior and outcomes of teachers is commonly held by students, parents and community. The non-educator is likely to believe that the role, actions and behavior of the "competent teacher" is always the same in any school or with any type of student. Thus, the student or parent may have an unstated image of the "ideal teacher" or on the other hand, be most concerned with associating any undesirable student performance with some deficiency in the act of teaching. While the random observations of students and parents may seem inaccurate or unjust, they do present evidence of attitudes, feelings and expectations of what a competent teacher should be. Systematic evaluation is not served well by summarizing the random observations that are reported. On the other hand, a survey of all parents, requesting their response to inquiries concerning desirable and undesirable behavior of teacher may provide a substantial body of information for the consideration of criteria of competence.

Of course the strong contention that must be drawn from several decades of research on teacher competence that is no single criterion has been found to be valid. uThus, it seems imperative to approach the issue of the evaluation of teacher competence as one that will most likely produce multiple criteria of competence. The criteria must also include process and product if they are to reflect the fidelity of implementing responsibilities assigned as well as the outcomes that are anticipated. There is no proven list of such multiple criteria of teacher competence but the following questions and criteria are frequently proposed.

- Does the teacher provide each student clear and unambiguous explanation of the content being studied?

- Does the teacher provide the students with exact and understandable assignments and expected student achievement?
- Does the teacher give clear explanation of how each student's behavior and achievement will be evaluated?
- Does the teacher demonstrate thorough understanding of the content and sequences of the curriculum for which the instruction is provided?
- Does the teacher demonstrate acceptance and understanding of individual student's development, needs, and differences?
- Does the student behavior and achievement reflect the teacher's ability to motivate students for constructive participation and high achievement?
- Does the attendance of students reflect a positive attitude toward participation in the class activities?
- Do the measures of student achievement show continuing progress from entry to exit from the class?
- Do students of different levels of achievement at entry to the class show comparable progress throughout the duration of the class?

. . . .etc., etc.

Such criteria as the aforementioned must be considered within the content of the educational system and the responsibilities that are given various priorities. The prioritization of needs becomes an initial step in determining the inservice interventions that are appropriate. The development of recognition and understanding of the roles and responsibilities of the teacher as defined by the system as well as the various outcomes that are expected by the students, parents, and system is an essential step in addressing the request -- "Why don't we identify incompetent teachers and do something about them?"

III. Which District Has an Instructional Program That Has Been Proven to Improve Test Scores?

"We members of the Board of Education are convinced that the local curriculum and instruction are responsible for the low achievement. Some of the district staff have suggested that we should budget for a comprehensive evaluation of our program and then develop new materials and programs to replace what isn't effective."

"This sounds like the same old request from the administrative staff to budget lots of money over several years to develop a new program. Instead of wasting time and money 'reinventing the wheel,' it would be more efficient to purchase a program that another district has proven to be effective for improving student achievement. Will you tell us where we can purchase such a program?"

This statement by the school board member reflects the frustration of the board of education with receiving continuing requests from the staff for additional time and money to develop new curricula, materials and instructional programs. The request for the identification of a program that was used in another district in which there were high test scores was probably based upon some assumptions that deserve examination. The request also reflects the unstated belief that there is a direct relationship between the books and instructional materials used and the test scores.

Myth or Assumption: Low student achievement test scores are due to the same conditions in all schools and all districts.

Fact: Low student achievement test scores may be due to a variety of factors.

Discussion: Many school districts use the same text books, study guides, and instructional techniques with a wide range of resulting achievement. An important variable in student achievement is the entry characteristics of the student population. Such student characteristics as environmental encouragement, or regard for academic success, or level of language development are commonly related to the academic proficiency that results from instruction. Academic progress has been found to be directly correlated with students' attitudes, skills and knowledge that have been accumulated in prior experiences both within and outside the school program.

Myth: The same materials and instructional techniques will be equally effective in all schools.

Fact: The effectiveness of instructional materials and techniques is dependent on a variety of conditions within the school.

Discussion: Instructional material or methods are often erroneously thought to have basic and generic power to promote student learning. This also reflects the belief that the teacher is less important for student achievement than the material or methods that are used. There is considerable evidence to demonstrate that the materials or techniques are similar to the tools that may be used in any trade: there are wide differences in the skill with which the tools are used and the speed and quality of the resulting product.

This assertion also disregards the time schedules, facilities and structures that are available for the implementation of instruction. In the "graded school" context there may be historical boundaries for the content or skills that may be presented to any grade or age. In a situation where high homogeneity of content is presented for mastery by all students there is low likelihood that a wide range of materials and contents would be acceptable and effectively utilized.

The developmental characteristics of the student population are critical variables in determining whether particular books or instructional materials will be effective for promoting learning. Student populations in various schools of a district may have large differences in general language development, study skills, motivation and general accrued academic abilities. Where there are wide differences in regions of a district or within schools there is strong likelihood that the textbook or materials that are producing high student achievement in one school will not be equally effective in another school with a different student population.

The variable effectiveness of any and all instructional materials across a broad spectrum of schools was documented by the extensive review of reading research sponsored by N.I.E. This review of more than twenty-five thousand reading research studies concluded that no particular material or reading program was clearly superior to others for attaining high reading achievement. These findings have been reinforced by very similar findings of the effectiveness of demonstration reading programs sponsored by the California State Department of Education, The Right to Read Projects, and the national evaluation of Title I programs. The studies all point to the conclusion that instructional materials or strategies of instruction do not possess a generic quality that will insure a desired level of student achievement in any school or with any population.

Myth: The same materials, techniques or instructional strategies will be equally effective when used by any teacher.

Fact: The effective use of instructional materials and strategies depends on many teacher-related factors.

Discussion: Instructional materials do not have indigenous qualities that assure common use and application. Moreover, the assertion that the critical variable is the material ignores the skills and knowledges that are usually required for effective use. It is also apparent that the assumption suggests that learning results from exposure to the materials rather than the dynamic interaction of the student, the teacher, the instructional materials and activities, and the context of the facilities and structures of the situation.

Experience with the often-recommended strategy for diagnostic - prescriptive - individualized instruction suggests that particular skills of the teacher are required and that these skills differ from those required by the lecture method or inquiry strategies. Various materials may require different preparation, presentation, or use, with students of various developmental characteristics.

Recognition of the different teacher-student interactive requirements may be observed in watching the implementation of such reading programs as "The Basal Series"; the Phono-Visual; the Sullivan; "The Language Experience Program"; etc. No materials are "teacher proof" as some would allege. Materials and programs are only tools that may be employed with varying effectiveness and success.

Myth: The purchase of a commercially published or district-developed program is less expensive than developing a new program in the local district.

Fact: The initial cost of the materials and instructional strategies is rarely the full expense of implementing a new program.

Discussion: The comparison of the cost of purchasing a developed program and associated materials with the cost for developing a local program over a period of several years will commonly suggest that the purchase of an externally developed program will be less expensive. The case studies described in Chapter 4 illustrate how these districts engaged their staffs in developments that took five or ten years and there is ongoing development and modification.

The critical issue in this question of expense is related to the often hidden costs of staff developmnt in understanding, acceptance and perfecting the skills needed to implement the new program. With the purchase of a program developed external to the district there is the need to educate the staff on the purposes, contents and procedures as well as making necessary adjustments in structures or schedules that will allow implementation of the program as it was designed. The full costs of these developments for staff and program structure may extend over several years after the original purchase of the program.

Thus, the comparison of expense of the purchase of an external program with a locally developed program must be based upon the full range of staff involvements that are required until the new program is operating and being consistently implemented throughout the district. This statement of accounting for the expenses required for implementation prompts the observation that it is essential that the district establish a procedure for monitoring the fidelity of implementation of a new program. The district is well served by a design for comprehensive program evaluation that will provide reliable information regarding the processes of instruction that are actually used as well as the outcomes that are measured.

Myth: Changing the instructional program is a simple, direct, quick and usually effective solution to the problem of low student achievement.

Fact: There are many difficulties and obstacles to changing the instructional program in a district and there are many possible advantages in changing programs.

The potential advantages as well as the difficulties are related to the characteristics of the community, the governing board of education, the administration and staff of the schools, and the student population.

Discussion: The enthusiasm for changing the instructional program is usually expressed by non-educators who assume that the materials and instructional techniques have been ineffective as evidenced by lower test scores than are desired. While some of the instructional staff may voice an interest in having new materials or new structures, there is a common tendency for instructional staff to rely on the techniques and procedures they have used and perceive as useful to manage the classroom and provide security for the day-to-day instruction. Frequently, the resident staff of a school has spent considerable time and effort in developing procedures that are related to evaluation of student achievement and the sequence for instruction from grade to grade in the school. Such developments are commonly grounded in staff preferences as well as being closely aligned with the personal skills of the teacher.

There is small likelihood that the purchase and adoption of new materials or programs by the administration of a district will result in immediate and effective use. Even if the staff may agree with the need for change to bring improved student achievement, a substantial period of staff familiarization, developing knowledge, and skills in the use of the new program will be required before full and effective implementation will be made.

The skill and enthusiasm of the staff in using any material or program is a strong predictor of the interest and success that students will exhibit with the use of the program. Students and teachers form important attitudes toward new procedures that are often related to the demands of the changing roles and interactions that are required by new materials or instructional strategies. The typical time required for staff familiarization with new procedures and changing revered practices which provided security lends much question to the assumption that an instructional program may be quickly changed through board adoption and administrative direction.

Experience with district efforts to purchase and install a new program based upon the recommendation of testimony by external authorities or other school staff or even by the single administrative edict has commonly produced the feeling in the instructional staff that they have been ignored. Instructional staff commonly hold strong beliefs that there is both science and art to teaching, and the teacher provides the interpretation and integration that enhances student learning. Most teachers also hold strong convictions that they are in the best position to make decisions regarding the materials and methods that will be most suitable and effective for the students in their school.

If an external program is praised as "proven to be effective" in the use by other districts, the local staff may be quick to reply that local students have different needs and characteristics than students in other districts. The local staff may also look for discrepancies between the new program content and what has been in use. Even though the differences may be small they will be used as evidence that the new program will not meet the needs of the student and thus cause further decline in test scores.

Any new procedure is commonly perceived as requiring "additional work" or "time consuming procedures" for management and use. The common reaction: "The new program demands so much paperwork and record-keeping that there is no time left for teaching!" Insofar as a new program requires skills and techniques that have not been used, the teacher may be threatened by the absence of techniques or activities that have been perceived as essential for the maintenance of discipline or management of the classroom. If the new program is perceived as confusing or disruptive to relations and classroom management that the teacher has felt to be successful, then there is high likelihood that the new program will not be fully implemented. The analyses of attempts to disseminate "model programs" under ESEA Title IV revealed that although a new program had been adopted and purchased by the centralized system for planning and management of instruction, less than 50 percent of the classes received the instruction that was prescribed by the new program. In some cases teachers made an initial effort to use the program but as there were unsuccessful experiences they reverted to the procedures previously used. It must be concluded that new materials and programs of instruction must be understood, accepted, and provide time for developing the skills for effective use if the new program is to enhance learning of students.

The potential advantages of changing the program of instruction are increased as there is recognition that an educational system is a "people system." Each school district serves a heterogeneous community and student population and the district staff provide an array of accumulated talents, skills, and abilities. There is a great need for using a process of adoption and implementation that involves the effected participants in the assessment of 1) the characteristics of the student population; 2) the relative effectiveness of the curriculum and instructional techniques that have been used; 3) the relative emphasis that has been given to various contents, skills and outcomes of the existing curriculum; 4) the skills and abilities and preferences of the staff; and 5) the relevance of the test performance that has suggested that there is need to change the program to improve student achievement.

Myth: Adopting an external program that has been used by other districts having high achievement test scores will improve the test scores of the district adopting the new program.

Fact: Test scores are influenced by many factors and one factor of central importance is the congruence of what is measured by the test that is used in a particular school or district and the content of the local curriculum.

Discussion: The various tests that may be used to measure student achievement have different congruence to the content, sequence, skills and knowledge that may be emphasized in a local program of instruction. In previous discussions in this chapter illustrations were mentioned of how various national, state or local tests may sample different contents or skills or may limit or extend the domain of knowledge that is used as evidence of mastery.

It is quite common for a school district to select a standardized test (from a variety of such tests) that the instructional staff and administration believe is most congruent and relevant to the content and expected outcomes of the local program of instruction. Changing the materials, and techniques of instruction may create a mis-match between what is measured on the tests used by the district and what is emphasized in the new program. Thus, changing the texts, instructional materials or strategies of the instructional program does not insure higher test scores as reported in another district. Test scores may improve, decline or stay the same depending on a variety of factors among which the congruence of what is measured by the test and what is emphasized in the program is of critical importance.

"After all, then . . . is it cost-effective to purchase a new program (either commercially prepared or developed in another district) that has been reported to have been used by districts with high achievement test scores?"

There is no correct answer for all school districts and for all conditions. The best answer for any particular school district wil be found by making a comprehensive evaluation of 1) the content or environment;

2) the financial resources available; 3) the beliefs and convictions of the leaders and opinion makers of the district; 4) the time available for staff involvement in development and inservice activities; 5) the time available prior to implementing a change; etc., etc.

The case studies illustrate that different administrative roles, different systems for making decisions and different roles of participation by the instructional staff may profoundly influence the nature and duration of change that is possible. At least two of the case studies described developments that required five to ten years of commitment of finances, staff participation and administrative priority for maintaining the effort and direction. Obviously, such conditions may not exist in many school districts; in the absense of the several essential conditions for local program development the answer may be -- purchase and adopt an externally developed program. The purchase of an externally developed program must always include the expectation that a substantial amount of time will be required to develop the understanding and skills needed for successful implementation of the new program of instruction.

The anticipation of immediate improvement of test scores of student achievement is ill-founded and will probably result in disappointment. Changes in student learning and performance on tests may only be meaningfully related to the instructional program as there is sufficient time to make longitudinal studies of student achievement on curriculum-relevant tests after the program of instruction has been documented as being implemented with fidelity and consistency.

IV. Why Don't We Ask The Superintendent To Develop A Comprehensive Plan To Improve Student Achievement And Present The Plan And The Related Costs Next Month?

This question often follows a series of suggestions and discussions of various conditions that are alleged to be responsible for unsatisfactory student achievement. Each suggestion is often stated as "the solution" with a request for additional funds. The board of education may be concerned that each request often refutes the importance of the other requests, and at the same time, wonder how the changes or additions may be incorporated into the existing program. Frustration with the variety of alleged solutions and the increased costs of the many suggestions may prompt the board members to propose that the superintendent develop a comprehensive plan that has a specific cost for each part of the plan and a total cost.

The persons requesting that the superintendent develop a plan and make a report in one month may very well have the belief that the superintendent has been putting off the request for actions to improve student achievement. Such a belief may be reflected by the citizen observation. . . "The superintendent just 'stonewalls' efforts to get a specific plan in operation, and continues to have different staff members submit proposals that deal with different parts of the program."

While this question/request for the superintendent to develop a comprehensive plan seems to be logical, it is made and based on several questionable assumptions.

Assumption 1: The superintendent knows or has access to the technical knowledge of the causes of unsatisfactory achievement of students as well as the interventions that will produce an effective and workable plan.

Discussion: The multiple causes of student learning or low achievement require a thorough analysis of both out-of-school conditions as well as the productivity of the existing instructional program. The request for the superintendent to develop a workable plan in one month makes the erroneous assumption that all technical knowledge is available and whatever plan is presented will be readily implemented within the existing operations of the system.

Assumption 2: The superintendent has major responsibility for improving pupil achievement.

Discussion: While the superintendent is responsible for the management of all aspects of the educational system, it is erroneous to assume that the superintendent is directly involved in the day-to-day implementation of the instructional program. In any educational system the instructional program is delegated to curriculum specialists, school administrators and teachers. Moreover, the day-to-day instruction by teachers may be required to be varied from class to class and school to school in response to the needs of the particular group of students. The variations in school populations, staff skills, knowledge and modes of implementing the curriculum demonstrate the improbability of a parsimonious plan to insure improved student achievement in the short span of one month. In reality, the responsibility for improving student achievement is shared by the entire staff who contribute important day-to-day implementation of the curriculum and the desired educational outcomes.

Assumption 3: The first action should be the development of a plan to improve student achievement.

Discussion: The request for an immediate development of a plan that will improve student achievement makes the erronous assumption that the "plan" will be universally effective with any and all students, and will be readily adopted and effectively implemented by all staff within the educational system.

Prior to the development of a plan, there is high necessity for careful problem identification, an assessment of the needs of the full range of diversity of the student population and the skills of the staff as well as the resources and facilities that are available within the system.

It is wise to consider the "Is - Should Be" model that proposes a documentation of what currently exists (What Is) and then make an equally detailed description of "What Should Be." The fairly common request for improved student achievement is a general and ambiguous request with little or no documentation of the priorities that should be assigned to the full spectrum of student development.

The careful documentation of the strengths and weaknesses currently existing within student achievement (as well as the strengths and needs of the educational system) provide an explicit basis for the educational staff, students, parents and board of education to consider priorities that should be addressed by a workable plan.

The impetuous demand for an immediate development of a plan prior to problem identification, needs assessment and establishment of commonly agreed upon priorities invites the probability of negative reactions. There is high likelihood that the "plan" will not address the various concerns of the public concealed in the general request for improved student achievements. There is equally high probability that the staff may view the "plan" as an unrealistic or impractical strategy in relation to what currently exists or what may be implemented within the resources and facilitites that are available. The development of a plan prior to the careful examination of the information gathered through the "Is - Should Be" assessment is a classic illustration of putting the "cart before the horse."

Assumption 4: Costs of a plan can be inventoried and assigned dollar values easily and quickly.

Discussion: The experience with attempts to implement a change in curriculum, special resources or instructional strategies suggest that even very simple changes have many time consuming and costly "start-up" operations.

A prime illustration of the difficulty of assigning time and costs to achieving a significant change in instruction is the complexity of the staff inservice required.

There is great wisdom in the statement -- it's a long distance from concept to construction. While it is relatively easy for one or several persons to conceptualize a plan for altering the instructional delivery system, it is very complex to achieve understanding by the staff who must not only understand the concept as well as the actions required to implement the plan. It is rare that an educational plan can be developed in such great detail that unforeseen needs will not emerge in the process of staff inservice and organizing the implementation. Such unforeseen needs and requirements may cause greater time and expense than specified by the initial plan.

Since the plan involves the mobilization of the actions of staff with diverse existing operations it is highly improbable to assign discrete dollar values to the stages of a statement of a plan to the final implementation.

Assumption 5: The statement of a plan for improving instruction and student achievement can predict and control the subsequent behaviors of teachers, administrators and other staff.

Discussion: The verbal description of a plan for improving student achievement may present an explicit series of changes that sound logical and possible. However, each change will require some changes in the behavior of the teachers and administrators responsible for the conduct of the program. As a plan calls for adopting new behavior there is a concomitant requirement that previous behavior be abandoned. It is questionable whether any plan (quickly developed without extensive participation by the staff) can predict how administrators and teachers will react to the plan or whether they will find it possible or feasible to implement within the other perceived demands of the school program.

The extensive efforts of the U.S.O.E. to disseminate "innovative plans and projects" to many school districts met unforeseen obstacles. While there was often initial acceptance of the need or rationale for the plan, when the local staff were engaged in implementing the plan a variety of conditions and established patterns of behavior were obstacles for carrying out the plan as stated.

A pertinent behavioral element that may confound a common plan for instruction is the preferred cognitive and communicative style of each staff member. The statement of a plan that may require a particular style such as didactic lectures, group discussion, student directed inquiry, etc., runs the risk of being alien to some individuals, or perceived as unworkable in the classroom.

Summary Point of View

The request for the Superintendent to develop a plan for the improvement of student achievement is commonly perceived by citizens or trustees as logical and appropriate. The request for a plan to be developed within one month shows insensitivity to the dynamics of an educational system and the amount of time required to conduct a meaningful needs assessment, problem identification, and prioritizing the multiple needs that are presented.

There is a compelling need to inform those requesting quick changes that the educational system involves the interactions of the content of the community with the students, the staff and the existing organization. The presentation of the variety of needs as well as the variety of perceived problems should lead to a "plan for planning" rather than the quick development of a plan.

The discussions of some of the assumptions that are made are not intended as rationalizations to ignore the request. In contrast, the examination of the assumptions suggests the compelling need to avoid hasty response and present a plan for planning that will allow ample participation and assessment of what currently exists as well as what should be.

The creation of a systematic plan for the development of a plan to improve student achievement affords an opportunity to have the staff develop understanding of the needs that are to be addressed by a modification of the existing program. The introduction of new materials, structures or instructional techniques without the understanding and acceptance of the implementing staff may produce hostility. The plan for planning allows discussion of natural resistance to change by instructional staff who may have strong conviction that the methods used are appropriate.

Since an educational system is a "people system," the importance of thorough communication of the needs that are cited as *requiring a change* as well as the alternatives that are suggested cannot be minimized. The absence of thorough communication between the administration and staff often results in the teachers believing that their knowledge and contributions are ignored. Certainly, a common reaction of the instructional staff to a quick edict that the instructional program will be changed is a passive attitude rather than full commitment to implement the new program as effectively as possible.

Skillful administrative leadership encourages planners, and supervisors of implementation to anticipate that some or many staff members will resist change. The communication with staff and community must patiently provide opportunity for expression of diverse points of view and then an equally patient but firm insistence that the alternatives must be prioritized and a decision for acceptance of a plan for improvement adopted. There is equal importance for common understanding that when the decision for a change is made there is a commitment that the plan will be utilized for a specified period of time to allow determination of the effectiveness of the plan.

CHAPTER 4

School District Stories

in which six districts tell how they
developed and implemented systems linking
testing, evaluation and instruction to
improve student learning . . .

CHAPTER 4

School District Stories

Introduction

Each of the following stories reveals hard working administrators and teachers dedicated to helping children learn. These six districts are "heroic" in that they persevered for long periods of time -- most for longer than five years -- to develop a comprehensive system for monitoring student learning. The systems they developed, however, were not simply descriptive. The tests they administered and the evaluations that they conducted were intended to point to remedies, and to modifications of classroom instruction. To the casual reader, the districts may seem very dissimilar to one another. So may the specifics of the work that was done by each of the central office staffs from start-up to the present. In this introduction, we will point out some of the similarities underlying the myriad of details. It is these similarities which may make it possible for other districts to recognize their own situation and to assess the extent to which they want to move in a parallel direction.

Environments. The six districts vary from small to medium size. With the exception of Vallejo and Cincinnati, they are primarily suburban, academically oriented districts with a relatively stable population. They have been stressed in the same way as other U.S. school districts: some population shifts; some budget problems; some external controls in the form of federal and state requirements accompanying federal and state monies. Nonetheless, when compared with big city districts, with rapidly changing conditions and large numbers of non-English speaking children, these districts seem to have had relatively more energy to turn to instructional matters. Good instruction was important in these districts, and to their parents, teachers, students, and administrators. The press for higher test scores -- believed by the community to more or less accurately reflect student achievement -- was accompanied by a district commitment to make itself the accountable agency for student learning.

In these districts, the federal and state role, not only in terms of providing money for special programs and imposing requirements for testing and evaluation, but also in stimulating new systems-oriented approaches to educational problems, should not be underestimated. The six districts represented here chafed, as did others, under paperwork burdens. Nonetheless, they often responded to the intent that the written regulations may have masked. San Juan's commitment to school level decision making was influenced by California's Early Childhood Program. Vallejo's interest in testing and its relation to instruction was stimulated by the California Assessment Programs. Cincinnati developed many of its needs assessment procedures and training outlines for local teams through federal funds. Newport-Mesa's work in their Objectives-Based Reading Program was encouraged by UCLA faculty who did much of their research with federal funds.

Initial district conditions. In each of the districts, you will be able to identify an "idea champion" and a stable core of staff. The idea champion or champions are individuals who were in possession of the care/

clout factor. They were in a position -- not necessarily the top administrative post -- where their caring could reverberate throughout the system. They had the clout -- whether formal or informal -- to translate their ideas into action. They had management skills. They could bring along, perhaps slowly at first, others with them.

The districts in which these idea champions worked supported their efforts in a number of ways. Persons with technical skills in developing or analyzing tests were either hired or procured as consultants. Academic consultants were often brought in to stimulate new thinking or work on a specific problem. Computer time was available from somewhere. Released time for teachers or summer salaries came out of district support, or was obtained from special project money.

Not only were there skilled leaders, willing followers, needed technical and managerial resources; there was also a vision of the possible and staying power to move in a direction in spite of setbacks. All the district stories describe false starts, mistakes, the need to start over. Los Alamitos, for one, purchased at least two testing systems before they moved to writing a curriculum scope and sequence. The Evaluation/Planner in San Juan felt isolated and lonely. He was a voice in the wilderness for a long time before gradually acquiring allies. Teacher resistance to his work, and often resistance from others in the central office was initially very high. But in all the districts, leaders kept their confidence in the testing, evaluation, instruction relationship intact even while modifying the specifics of what they were doing.

Developmental sequence. None of the districts describes a planning process that was formulated at the beginning and served as a framework throughout the development cycle. Rather, things seemed to emerge, evolve, and get made up as the districts went along. One thing led to another.

Each district began its process at a different place. Newport-Mesa first examined what the community expected from the school districts. Los Alamitos began by purchasing tests. San Juan wanted to encourage and empower school site planners. Vallejo and Clark County both reacted to low test scores by trying to express a district philosophy and direction. Cincinnati reacted to negative community opinion and slipping budgets by encouraging neighborhood-business-school relationships.

Each district proceeded, also, through a series of stages, coordinating their operations in different ways as they went. Newport-Mesa and Los Alamitos now have a testing system referenced to objectives contained in a district-wide scope and sequence. Los Alamitos, especially, has tied this framework to a particular diagnostic/prescriptive teaching strategy, supported by indexed media and materials. Clark County also mandates particular teaching elements in the classroom and couples this with emphasis on principal supervision of teachers and central office supervision of teachers.

San Juan and Cincinnati, on the other hand, are decentralized in their use of data from tests and surveys. Both collect information about parent, student, teacher perceptions and organize them into easy-to-read packages for school groups to use. In both these districts, the central offices provide training support and guidance for local teams who work through an annual proces to allocate school resources to meet high priority school needs. Vallejo is doing something on a number of fronts. Having started with school-site analysis of instruction in relation to the CAP tests, they moved to strong staff development programs, and finally to developing a district-wide scope and sequence.

Role of major actors. The district central offices have a strong role to play, but the role differs somewhat from district to district. In all districts, a variety of tests are required. District staff analyzes or has analyzed, the tests. District offices provide staff development to teachers; or in the case of San Juan and Cincinnati, leadership training to citizens. District offices also provide support or guidance to teachers in different ways. In Los Alamitos, texts and other materials are cross-referenced to the tests. A learning specialist works with teachers in classroom management problems. In Clark County, the entire supervision and management system is set up to ensure accountability for clearly defined roles and responsibilities.

The teachers' role in development of the process differed from district to district. However, it seems that, in all districts, there was initial resistance, followed by gradual commitment as teachers themselves worked on the system. Nonetheless, it is not clear from the accounts whether all teachers now believe, once the system is in place, that instructional improvement commensurate with central office efforts has resulted.

Impact. The district stories imply more than they prove. With the exception of Clark County, there is little evidence presented of a dramatic increase in student scores. This, however, may not be the only evidence of impact that we should consider. It does seem clear that, for many districts, their public image has improved. Los Alamitos, Clark County, San Juan, among others, clearly enjoy reputations as forward-looking, well-managed, self-renewing organizations. Teachers in all six districts as well as administrators may have a feeling of professional well-being different from that in most districts. The districts all have a clear sense of purpose and of priorities.

[More to come.]

LOS ALAMITOS' EXPERIENCE:
ITS CRITERION-REFERENCED TESTING SYSTEM

by David Hatton

The Los Alamitos Unified School District is a small school district now consisting of five elementary schools, two middle schools, and one high school.

Our district had its beginning in 1881 as a one-room school with an enrollment of 43 students and a first year budget of $700. It remained a rural one-school district until the housing boom of the mid-1950's, when our name was changes to Los Alamitos School District to provide community identity. The district continued to grow until it reached a maximum of seven elementary schools with an enrollment of 4,000 in the early 1970's. As has been the case in many districts, declining enrollment since then has resulted in the closing of two schools.

In June of 1979 residents of the Los Alamitos School District voted to form a unified K-12 district which became effective in July, 1979. The elementary district, therefore, ceased to exist on June 30, 1980, concluding 99 years of operation. The Unified District, with an enrollment of 5,200, is comprised of the former Los Alamitos Elementary District and a portion of the Anaheim Union High School District (Oak and Pine Junior High Schools, and Los Alamitos High School).

The Los Alamitos Unified School District is an upper-middle class suburban setting. The large percentage of parents who are college graduates have very high educational expectancies for their children. As a result, the district has been a leader in the development of innovative educational programs, and has consistently scored among the highest in the State on achievement tests, while operating at or below the State's average per-pupil expenditure.

Because our parents emphasize the importance of academics, there has been a continual demand from the community for student progress reporting. Our parents want to see consistent student progress; progress which they expect to be documented in a clear and understandable fashion.

Development of the Testing System

Formative Years

In 1970, the district attempted to meet the need for student progress reporting by purchasing test items to measure a sequence of behavioral objectives. The objectives turned out to be improperly sequenced. Teachers hated them and eventually district administration discarded them.

But there remained the need for student progress reporting. However, an equally compelling need began to emerge -- individualization of instruction. Though the behavioral objectives system had not worked, district administration realized that a properly developed test could not only satisfy the parent demand for progress reporting, but also be a vehicle for providing instruction appropriate to a student's current level of achievement. Whole-class instruction could be replaced by small group instruction. Students could be grouped according to their performance on the test and instructed according to their current performance levels.

Rather than purchase an available testing system, district administration formed teacher committees to write their own tests. The feeling was that tests written by district's teachers would be better received by other teachers than tests developed by an outside agency.

In 1973, the first district-developed tests, known as PAL (Pupil Assessment Lab), were administered to students. They were intended to monitor student learning so as to feed back information to teachers about how individualized instruction was working. They also were seen as a way of reassuring parents who were concerned about their children's progress.

Teacher reaction to PAL was extremely negative. Everyone complained. Teachers today remember their complaints:

- It was a waste of time. It didn't tell us anything we needed to know.
- The children were confused by the tests.
- We didn't know what to do with the results.
- We didn't want to be judged on the basis of what our children did on those tests.
- It was not coordinated with anything we taught.

Teachers made their complaints known to parents and to the Board. It was a hard time for central office staff. We wanted the testing system to work.

We appointed teacher committees to try to revise the items. We responded to the need for coordinating testing and curriculum by beginning to work on a district level instructional continuum. A group of volunteer teachers were paid to work on a reading continuum during the summer. Their work continued through the school year.

By 1976, teacher committees had generated a reading continuum (RIC) and the beginning of a math continuum (MIC). By 1976, the PAL criterion-referenced testing system had been scrapped. Teacher committees had written reading, math and language arts test items to form their own Criterion-Referenced Tests (CRT). The tests were developed by the district's teachers and designed to support district-adopted instructional continua. A more detailed description of our CRT's comes in a later section.

Staff Development Program

Realizing that a testing program by itself could not improve instructional practices, we at the central office instituted a staff development program in the early 1970's. Convinced that a staff development program was at the very core of a support system for the criterion-referenced testing program, we hired a full-time staff development coordinator in 1976. Our program was designed to provide teachers with the skills necessary to act upon the results of the district-developed testing system.

Currently, the staff development program includes two levels of instruction. Level one is based upon the Madeline Hunter orientation to teaching. Nearly all teachers participate in level one. A modified version is provided for substitutes and aides. Covered in level one are teaching techniques, including task analysis, diagnosis, prescription, instruction and evaluation. Teachers attend the staff development classes on release time and then practice the techniques in their own classroom. Follow-up visits are made by the staff development coordinator.

Level two programs are of two kinds. The first kind is an extension of the level one activities with attention to individualizing instruction by choosing among alternatives. Additional sessions cover matters such as diagnosing and developing teaching strategies for meeting affective needs, using inquiry skills techniques, and teaching specific curriculum content such as reading, writing, or mathematics.

The Learning Specialist

At the same time, the staff development program was placed in operation. We in the district office decided to give the responsibility for coordinating the newly developed testing system to a learning specialist assigned to each of the schools. We regard the learning specialist as a superteacher, not as an administrator. Approximately 40 percent of the time of the learning specialist is spent working directly with students. The remaining time is devoted to a range of other activities. The learning specialist works with teachers in their classrooms to provide instruction meeting the needs of students as indicated by the test results and by teacher observation. The learning specialist sometimes conducts demonstration lessons and makes suggestions for new materials. The learning specialist meets with teachers when the test scores come back and helps plan for the next several months. A small portion of the learning specialist's time is devoted to in-school inservice activities. The learning specialist is also responsible for diagnosing new students as they come into the classroom and facilitating the placement of children in instructional groups.

Media Center

Within every school in the district is a materials and media center staffed by a full or part-time specialist. The district essentially orders materials that teachers and learning specialists request. The central office coordinator of materials each year asks the teachers what they want, looks at areas in which materials need to be supplemented, and meets with the media specialist to facilitate the acquisition of equipment and supplementary curriculum materials. The learning specialist in each school often works closely with the media specialist to try out new materials as they arise to determine how appropriate they are for particular types of children. The learning specialist may try out sample materials sent to the schools to decide whether they are suitable for ordering. The media specialist in each school often functions in an advisory capacity to the learning specialist and to the teachers, recommending books, films, adjunct materials which he/she thinks appropriate.

Description of the CRT

Originally developed solely as a measure of student progress, the tests initiated in the early 70's have evolved into a system which assures a linkage between instruction and test results. Supported by the staff development program and by other support personnel such as the learning specialist and the media specialist, the system currently is in full operation in grades K-8. Highlights of our testing system:

- All tests are directly linked to a kindergarten through eighth grade instructional continuum for both reading/language arts and mathematics.
- All tests are computer-scored. Results are formatted in easy-to-read computer reports to teachers, parents, and site and district administrators.
- Information about the testing system and how it works is taught as a part of the staff development program.
- The tests are reviewed on an ongoing basis with revisions occurring four times over the last several years.
- The testing program has been recognized by state and national teachers, administrators and college educators as an exemplary model appropriate for many school systems.

The costs of implementing a system like the one currently operating within Los Alamitos Unified School District are high, both in terms of dollars and in terms of teacher and administrator time and effort. But the benefits are also great. Teachers can use the tests to group students according to instructional needs. Teachers and site administrators can use summarized test results to identify school-wide instructional strengths and weaknesses. This information can be used to set goals for the coming school year. District administrators can use the results in the same fashion to identify appropriate instruction program priorities for both the immediate and distant future. Finally, the tests can be used to meet continuing mandates from the state and federal government for the setting of local graduation standards.

Test Uses

Perhaps the most important use of the testing information is made by the classroom teacher in planning for the instruction of his/her students. The test is referenced to a graded sequence of instructional continua for reading/language arts and math, thus enabling the teacher to determine exactly what skills have been mastered by each student and which still need to be learned. An example may serve to illustrate this important use of the test. Assuming a first grade teacher plans to teach calendar skills to his/her class. The teacher would first consult the district instructional sequence for teaching calendar skills. A section of the sequence is reproduced below.

Task Analysis	Grade Level
1. States the days of the week in order	K-1
*2. Identifies the number of days in a week	K-1
3. Locates the names of the days of the week on a calendar	1
4. Reads numbers to 31	1
5. Locates a given number on a calendar	1
*6. Identifies days of the week for a given date	1-2
7. Identifies the number of dates in a month for a given day of the week	1-2
*8. Identifies the dates of a given day of the week	

The learnings which have asterisks are known as benchmarks and are tested. After becoming familiar with this task analysis, the teacher administers benchmark tests as appropriate for each student. Generally the tests are given to small groups of children. Individual performance on the test will enable the teacher to reorganize student groupings according to level of mastery of skills along the continuum. Not only can the teacher group the students for instruction based on their current level of achievement, but the teacher can also project student growth over the course of the school year. Those students not able yet to identify the number of days in the week might be projected to locate a given number on a calendar by the end of the school year. But a student who already is able to identify the dates of a given day of the week might be projected to master one of the learnings that falls at the second or third grade level of the task analysis.

It's important to note that the testing is designed as a gross diagnosis of student performance. After teachers review the test results, they decide whether to move back or forward along the continuum and develop their own form of assessment as they proceed along the instructional sequence.

In summary, individual teachers use the test for gross diagnosis of student's current level of performance and for projecting and measuring student growth over the course of a school year. Teachers as a group also work with school principals to identify instructional priorities according to student performance on the tests. Since the tests measure several content areas within reading, language arts and mathematics, it is possible to identify student strengths and weaknesses within each of these disciplines. As an example, analysis of global performance in reading might reveal that as a group, seventh graders had strength in identifying the main idea of a short story, but that they had difficulty in predicting the outcome of the story. The teachers and their principals might identify "predicting outcomes" as a high priority instructional area for the coming school year.

The testing system also allows teachers and administrators to determine the percentage of students performing below, at, and above grade level. This information may be used for a variety of purposes, including the validation of other standardized tests mandated by federal and state regulations.

Another use of the testing information occurs at the district level. District administrators can review test results with site administrators to set district and site-level instructional priorities.

Finally, the testing system can be used to meet proficiency standard requirements currently enacted in some states. California State law requires all school districts to establish high school graduation standards and to provide a warning system beginning in fourth grade for those students not making satisfactory progress toward graduation. The testing system developed in Los Alamitos fits in nicely with these requirements. Los Alamitos students who are found to be two years below grade level in any of the tested areas are identified as needing additional remedial assistance. In this way, the system is used not only to guide instruction for all students, but also to identify those students who need remedial assistance.

Factors Contributing to Successful Implementation

The system sounds attractive. But I've already indicated that it was costly to develop, and that may make it not feasible for all school organizations. A number of factors expedited the implementation of our system in Los Alamitos:

Consistent direction from leadership. Though school board composition and superintendents changed over the last fifteen years, commitment on the part of these leaders was strong. Since the system is expensive, support from leadership within the district must be strong and consistent.

Principals as instructional leaders. Principals are viewed as and are trained to be, the instructional leaders at their schools. Because of this, administrators are knowledgeable with regard to curriculum and instruction in each classroom. Teachers and parents know that principals' administrative responsibility is not limited to public relations, scheduling and discipline. Principals know about instruction and can communicate knowledgeably with teachers about the learning of students within each classroom.

Staff development. Our commitment to ongoing staff development has expedited the implementation of the testing system. Through staff development programs, teachers learn about the tests but more importantly about the instructional continua the tests are designed to support. Teachers not only learn how to diagnose appropriate learnings for a particular instructional group but also experience new and effective instructional techniques. The emphasis in the staff development program is upon continuing education for teachers. Our focus is not upon teaching teachers to test. Rather, it is upon teaching teachers to teach. Testing becomes a minor though essential part of instruction.

The widespread participation by teachers in the staff development program has had the consequence of providing teachers with methods of acting upon the results of the district-developed tests. Without this instruction, teachers would not have the tools to effectively use the results of the tests and non-use or inappropriate use would follow.

Adequate funding. The development of the testing system was funded at an adequate level by our district. We also recognized from the beginning that implementation of the system would be expensive. Not only was appropriate computing equipment purchased to implement the system, but well-trained personnel were also employed to operate the equipment.

The Learning Specialist. In order to successfully operate a testing system similar to the one operating in Los Alamitos, someone must serve in the role of learning specialist at each school site. The learning specialist is a resource to teachers and assists the teacher in acting upon prescriptions implied by the student's test responses. The learning specialist brings the teachers suggestions and possibilities for instructional alternatives; more importantly, he or she provides extra instructional time for students who need it. The services provided by the learning specialist are facilitative. He/she is seen as a resource for teacher, not as an intruder in the classroom.

Recruitment and selection. The final, yet the most important component to the successful implementation of the testing system, is the expertise of the teachers within the district. The district has long held to a policy of rigorous personnel selection. The selection process is comprehensive and requires candidates to demonstrate their teaching skills prior to being selected. As a result of this rigorous selection process, teachers within the district may be generally characterized as enthusiastic and committed to teaching. They set high expectations for themselves and for their students. The district's teachers definitely are its most important resource.

Mistakes To Be Avoided

The road to full implementation of our testing/instructional program was not easy. Mistakes were made along the way. A few that would seem to apply to others interested embarking upon a similar, but hopefully less rocky path, are as follows:

Don't let test development precede curriculum development. In 1973, when PAL was developed, Los Alamitos teachers and administrators selected "performance outcomes" and then wrote test items to support them. The performance outcomes were not selected from an instructional sequence and therefore were meaningless in terms of informing the teacher where a student had been and where he/she now needed to go. The test based solely on performance outcomes became something to be passed and then forgotten. Much like the high school proficiency exams, which are required in some states, much effort went into preparing students to pass the exam, but once passed, teachers had no direction or incentive to move to higher levels. It's important to note that not all of the learnings in our instructional continuum are selected for the test. If the test is failed, the teacher knows he/she must move back along the continuum to find the point at which to instruct the student. Passage of a test on a learning item means the teacher can look to the next step in the instructional sequence. Obviously, the instructional sequence must have been carefully developed and be subject to revision by teachers.

Don't overemphasize test results. Tests are readily over-interpreted by teachers, administrators and parents. Strangely, those teachers who seem to be most threatened by tests, especially tests given to every student in the district, are the same teachers most apt to over-rely on test results as measures of student progress. While threatening to them, test results seem to offer a certain sense of security to anxious teachers. A teacher lacking confidence in his/her teaching skills may emphasize test performance as a substitute for good teaching. We know that good test results may not always indicate good teaching practices. Conversely, good instruction always leads to good test results.

Administrators and parents sometimes contribute to the overemphasis on test results. Tests have an "objective" aura about them. However, if results are not combined with such powerful subjective measures as observation of classroom climate, student oral work, and student interviews, test results are prone to inflation. A balanced evaluation of student progress by parents and administrators as well as teachers is essential to the effective use of a testing program. At least annually, our administrative staff meets as a group to discuss strategies to avoid overemphasis upon testing. Generally, such strategies include highlighting other measures of overall student performance which indicate strengths and weaknesses in a particular school's instructional programming. Such brainstorming by administrators and teachers of alternative measures of evaluating instructional programs has been quite productive.

Just as we must reassure teachers who are resistant to testing that the test was designed as a support to instruction, so must we encourage teachers who are overly test conscious to use test results as only one measure of student progress. In the latter case, teachers may become so preoccupied with "good" test results that students are pushed to higher levels measured by the test at the expense of broadening or enrichment of skills at a lower grade level that are important but not tested. The very best teachers are often trapped by this problem, especially because of the pressure for high test scores from our parent community.

The school principal is one person largely responsible for controlling the overextension of testing. Unless the principal reinforces other measures of student progress, students will be pushed to higher academic levels without receiving important supplementary instruction. Striking the balance between broadening of current skills and extension to higher level skills is an important responsibility of each site administrator.

Don't rely on outside agencies to develop your testing system. In 1970, our district contracted with a testing company to develop a test and to process the testing information. While the test items were good, the data processing was slow as is often the case when an outside agency is involved. The turn-around of testing information was so slow as to make the results meaningless to our classroom teachers.

But even if the data processing had been good, there still would have remained a need for teacher "ownership" of the test. Such ownership comes about only when teachers within the district have actually participated in the test development process. While it probably is more costly to take the time to train teachers in item writing skills and to do your own statistical item analysis, the long-term benefits gained from teacher identification with the test will go a long way to insure successful implementation of the test.

Questions Commonly Asked of Us

Throughout the years of implementation of the district's testing program, a number of questions/concerns have arisen again and again. The most salient are summarized in response to the following questions:

1. Where does teacher accountability fit?

Try as we in the central office may to avoid the issue of teacher accountability and test results, the issue comes up every time a new teacher learns about the test or an experienced teacher complains about overemphasis of testing. This issue should not be avoided. In fact, it should be pointed out that an overemphasis on accountability leads to inflated and meaningless test results. Certainly, accountability as to be a part of a testing program. Indeed, the source of the pressure for accountability should reside with a teacher's desire to move his/her students toward comfortable and attainable goals. As stated earlier, overemphasis on test results by teachers, parents, or administrators usually leads to inflated test results. A tool which can be extremely helpful to instructional improvement is thereby rendered meaningless.

2. Would you advise developing our own system or purchasing someone else's?

Develop your own system, if possible. The advantage of having those who have to implement the system, namely the teachers, identify with the system, far outweigh the advantages of developing the system outside the district or purchasing another system developed by someone else. Even if some of the technical aspects of test development have to be sacrificed, it is much better to have a system in which teachers are confident.

3. How about staff development?

This program is essential. Without it, all emphasis will be upon testing, rather than teaching. Again, the object of a good testing system is not good test results in and of themselves. The object is to support and supplement good teaching practices. An invaluable aid to emphasis upon instruction rather than testing is the presence of a staff development program. Teachers must be given the opportunity to refresh and update their teaching skills if good instruction is to be offered within their classrooms.

4. How much will it cost if we do our own data processing?

A lot! Our computer programming costs alone have been $30,000 over the past seven years. The equipment to operate the system has cost more, though the computing equipment is used for a variety of other purposes, including business management, and student attendance and grading. Sufficient teacher involvement is costly too. Teacher committees met once monthly for six hours a day over a two-year period to develop the reading/language arts and math tests currently used in the district.

5. How do you think testing should be implemented?

Slowly! Many districts have made the mistake of moving too quickly with the implementation of a testing sytem. Simply put, the users of the tests -- students, teachers, and administrators -- weren't ready. Without the involvement of these groups, the tests are doomed to failure. Start with a needs assessment of the communities the test will serve. If only one school is interested, implement the system on a trial basis there. Most importantly, make sure the school board is committed to a venture which will be time-consuming and costly, but which will lead to the rich benefits of improved student progress.

6. What differences occur at the elementary and secondary level?

As most educators know, the orientation of elementary and secondary teachers differs markedly. Typically, secondary teachers see as many as 150 students a day. Students are seen in groups of 30 for slightly less than an hour. Elementary teachers generally work with 30 students for the entire school day. The large numbers and relatively small amounts of time with students tends to make secondary teachers resistant to any testing system requiring extensive amounts of time to administer and score. Since the system described in this paper does require a significant commitment of time on the part of the secondary teacher, release time was given each high school teacher for the purpose of scoring the tests and interpreting the results with the learning specialist. This released time paid rich dividends in increasing teacher morale towards the new testing system.

As our CRT system was expanded to the seventh and eighth grades, the question of standardization of instruction arose. This was not a new issue. A similar concern was expressed when we developed our instructional continua at the elementary grades. The concern stems from the fact that it is necessary to have a standard curriculum if a standard test is to be developed. Some teachers indicated that their freedom to select materials and course content was threatened by the need for such a standardized continuum. Our response to all grade levels has been that the purpose of a standardized continuum is to assist student learning of skills that have been deemed important by teachers in the district. The continua were developed by teachers for teachers; the intent was to assist instruction rather than control it.

In no way is the "how" of instruction dictated by the continuum. Teachers are given full responsibility for selecting and implementing instructional techniques to teach the skills listed in the district continuum.

Summary

The system described in this paper has the support of the teachers, parents, and administrators in the Los Alamitos community. Its greatest asset is that it was developed by the district rather than by some outside agency. It works because test results are not overemphasized and because teachers who use the system are provided with the training necessary to use it effectively. Provision of proper training has meant the addition of support personnel in the district. The bottom line is that having a system similar to the one described in this paper is much more expensive than other testing alternatives. However, the expense is far outweighed by the excellence of instruction that is made possible by such a system. School entities who find themselves in a situation similar to that of Los Alamitos and who can afford the cost of a district-developed system should certainly consider this alternative.

THE NEWPORT-MESA EXPERIENCE: THE HISTORY OF A COMMITMENT TO CBE

by Dale Woolley
and Nola Rochelle

Many "old-timers" in this district recall an incident that occurred to Don Hout. Nearly fifteen years ago, Don was an elementary principal who was paid a visit by an interested father during the summer. The father was enrolling his fourth grade youngster for the fall semester and was interested in helping his son make a successful transition from his former school district to Newport-Mesa. He asked Don, "What is it you expect a typical fourth grade student in your school district to know when he or she enters fourth grade? What skills do you expect him/her to have already mastered?" Don was concerned that this very reasonable question was difficult, if not impossible to answer. He could talk in terms of curricular materials to be studied, but not specific instructional goals to be mastered. This type of inquiry from the middle to upper income parents that characterize the Newport-Mesa District leads one to understand why in 1965 the Board of Education selected Leland Newcomer for its superintendent.

Lee Newcomer met the Board of Education's criteria for a new leader. The board was interested in finding out what the main educational priorities of a school district should be. Lee Newcomer was schooled in accountability and suggested a data gathering system that would provide them with just such information. To accomplish his mission, the new superintendent brought on board a new director of research and development from his native Clark County, Nevada. The new driving force was to be Les Shuck, an administrator who would provide a detailed blueprint for a CBE system that would guide

Newport-Mesa for ten years. For most of this time the district would be steered by three men -- Lee Necomer, Norman Loats (then Associate Superintendent), and Les Shuck -- whose joint philosophy it was that the district should be in control of the "what" of education (the important skills to be learned), while the schools would be in charge of the "how." For many years, the district was to be decentralized. Each school would have its own budget and the freedom to achieve the "what" in whatever way the principal and staff saw fit.

The project which initiated Newport-Mesa's question for the "what" of education was called the Instructional Tasks Project. The I.T.P. was written by Dr. Shuck with the help of some educational consultants and funded by ESEA Title III or federal monies. Its goal was to determine the contributions that the community expected from the school district. In process, it involved in-depth interviews with a stratified random sample of community members. Participants were asked to describe specific incidents from their past where the behavior of young people served as good or bad examples of the kinds of skills and knowledge they thought young people should have. The results of this one and one-half year study included the finding that some skills and knowledge were considered more important than others. Exactly what primary and secondary responsibilities the schools should assume became the Board's most important policy. It was adopted in 1970 and entitled the "Statement of Educational Principles," thereafter known as SEP for short.

SEP dictated that the schools' primary responsibility was to help students develop their maximum intellectual capacity in twelve specific skill and knowledge areas. These twelve areas included:

reading
composition
listening
speaking
computation
mathematical systems
science
social studies
fine and practical arts
doing
thinking
language systems

Since reading is always an area of great concern, the district began with this subject area in its first attempt to further specify the skills comprising this general skill. As it would on many occasions, the district would turn to university consultants and research and development centers for suggestions as to how to proceed. The project of reading led to OBER.

Although the development of behavioral objectives was becoming quite a popular thing to do in many education circles, the district had a hard time finding examples where an entire curricular area had been translated into a continuum of specific instructional outcomes. The new director of the district's Development Lab, Bob Otto, hoped to adopt or adapt the work that had been accomplished outside the district to the district's needs. Otto found that Professors James Popham and Rod Skager were developing a continuum of instructional objectives in reading, grades 1-6. Newport-Mesa then entered a joint venture with U.C.L.A's Center for the Study of Evaluation, where ths continuum was being put together, that allowed the district to use the objectives and test items., This continuum was to be called OBER for Objective-based Evaluation--Reading and was to represent not just the objectives that were important for every student to master, but rather all the objectives that could be used to define what was meant by "reading" at the elementary school level. Interested teachers in the Newport-Mesa schools were asked to review the objectives and choose those they thought important for students to master at various grade levels. Most participating teachers did this on school time and for no extra pay. Objectives chosen by participants from more than half of the schools were then used to conduct district-wide assessments. Otto and his group could then paint a picture of how students were performing on these chosen objectives. The results of this assessment yielded more than the usual data.

One unexpected piece of information was not the least academic in nature. It was learned that school staff members have a strong tendency to be unconcerned about the process of choosing instructional outcomes until assessment begins to take place in their classrooms. The OBER assessments generated teacher and administrator comments that the district was trying to "mechanize" the creative process of education. Others said that "you just can't measure what is taught in my classroom (or school) in specific instructional outcomes!" Despite such commentary, Dr. Shuck, Dr. Loats, and the new superintendent, Bill Cunningham, forged ahead, using OBER as a model and learning its lessons on staff development.

Since no other complete system such as OBER was available, the district set out on its own to create needed continuums in other subject areas. With some consultant assistance -- notably Dick Harsh from E.T.S. and Dr. John McNeil from U.C.L.A -- and the district coordination of Cora Schultz, newly-appointed resource teacher for the Development Lab, the district set up twelve teacher committees. There was one committee for each subject area in the SEP list or priority subject areas. Each committee was charged with the duty of developing a continuum of objectives. Through a competitive interview process a capable teacher was selected as chairperson for each committee. S/he had the freedom to select the rest of the committee members. Each was paid for 22 extra days of work per year. Frequently, the chairperson and his/her committee members were provided with substitutes for all-day organizational or work meetings. Each committee had a small budget in order to bring in outside consultants to help them when necessary.

As the SEP committees began to work writing behavior objectives to define each subject area, twelve more committees were created to develop test items to match the objectives being developed by the SEP groups. These committees were concerned with "program outcome evaluation" and were known as the POE committees. Interaction between the SEP and POE groups resulted in twelve continuums of behavioral objectives known as "universes" -- tomes that were to a fledgling CBE program what the English dictionary is to a foreign-born speaker of our language. This interaction also resulted in banks of items correlated to these objectives, and to many heated discussions between the two committees about the merits of certain objectives. The writing of test items, it was discovered, caused the rejection of many an objective as unclear or impossible to measure. Test items served to test the objectives themselves as we went along.

At this point Research and Development and Development Lab staffs were beginning to anticipate some district-wide assessments. Items were available and the time was drawing nigh. The missing link was the need for a wider teacher exposure to the work accomplished and for a thorough surveying of teachers to determine the grade level where each skill would be considered appropriate. Universes of objectives were sent out to each school and the process known as U.A.L. was begun. U.A.L. stood for Understanding, Acceptance, and Level of Assessment. Each school was required to complete the U.A.L. process on a set of objectives sent to them. In most cases the requirement was satisfied by one teacher, a small group of teachers, or the school principal completing the task. The U.A.L. process worked best when teachers did participate in the process. It was later discovered that several objectives that were "accepted" and "leveled" by principals were objectives which teachers singled out as unworkable and inappropriately leveled. Once again, a test instrument "got everyone's attention" and exposed what seemed like a harmless shortcut arising from the usual indifference. The U.A.L. process did serve to make schools more aware and more educated in the growing competency-based system as it also gave the SEP committees more revision to do and the district some notion of what objectives to assess.

The long-awaited SEP assessments naturally followed. With the assistance of Dale Woolley, Director of Pupil Personnel Services, a schedule was developed so that a given SEP subject area would be tested district-wide every two years. A matrix sampling approach was devised with item sampling used on a 100 percent student sample. First assessments included reading, computation, and math systems. Following assessments included language systems and composition. Each year additional areas were tested. By 1977-78 all SEP areas had been tested except for science, fine and practical arts, and doing. Teachers responded to these assessments in several ways. Those involved with the SEP or POE committees were enthusiastic and filled with anticipation. Many teachers at school sites were concerned that results not be used to evaluate their effectiveness in the classroom. During this time of initial impact it was fortunate that due to matrix sampling the district could not report results on individual students or classroom units. Only grade level and school scores could be made available. After several years the SEP assessments had run their course and, due to budgetary constraints, were discontinued in 1979.

What took its place was called Student Progress Monitoring or SPM. The new computer program written by data processing consultant Hal Roach, was to permit the very individual student results that many teachers hoped for and that others feared. SEP assessments had been thought unnecessary by many teachers. They often objected to the expense of such testing and saw no "relevance" for the classroom teacher since they did not personally choose the objectives on which their students were tested nor did they receive classroom level results. The new SPM program was to solve these problems and quiet many of the objections.

SPM began as a pilot project at Paularino Elementary School where an interested, forward-thinking principal -- Bill Knight -- asked for district funds to develop a small, individual student testing system with computer support. Based on his positive experience, a district-wide committee of teachers and principals agreed to meet with central office staff to describe the type of computer programs desired. This program proved to have the following advantages:

(1) Individual teachers of groups of teachers were able to select specific objectives from the various SEP Universes, then develop tests using items from the district item banks and administer these tests to their students.

(2) The district would then machine-score the resulting answer sheets and provide reports back to schools indicating which students had mastered which objectives.

(3) Teachers also had the option of writing their own objectives and items and assessing students with a totally "home-grown" test instrument. Although this option was not selected by a large number of teachers, it did give some teachers the feeling that they were not bound to use only the objectives that were in the district banks.

Although the SEP district assessment program was mandated during its duration, SPM started out as an optional testing program that individual teachers and groups of teachers were encouraged to use. Schools that were not performing well on the district SEP assessments were encouraged by central office staff to use SPM to pinpoint those students who could benefit from added instruction on the particular outcomes in question. The introduction of SPM into the district was done on a low-key basis. Its use was voluntary and increased each year after its introduction. When the SEP district assessment was suspended in 1979, SPM, along with the newly developed Minimum Graduation Proficiency Testing Program, became the main assessment tool of the district's competency-based education program.

In 1976, California passed a piece of legislation called the Hart Bill which required the establishment of minimum graduation proficiencies in the basic skill subjects of reading, computation, and writing. The first class to be affected would be the class of 1981. Although many California districts were caught unaware and unprepared, Newport-Mesa was indeed ready for the challenge. Having complete "universes" of objectives in many subject areas including the basic skills, the district merely gathered community leaders, parents, administrators, SEP and POE chairpersons, teachers and some students to select important objectives for consideration as graduation proficiencies. Although such a process "worked" to a certain extent, a rather positive-thinking and ambitious group was turned loose on a task that somewhat resembled the supermarket game show where contestants are given a market basket to legally "loot" all the merchandise they can in a specified period of time. It had all the appeal and results of the proverbial excursion by the "kid in the candy store." After the first series of lengthy meetings, the committee had identified over 1,200 objectives for further study as possible proficiencies! As would be expected, this unwieldy list of skills was finally reduced to a more reasonable number while still being comprehensive of the curriculum areas to be tested. The final versions of the test were at last created by committees of secondary level department chairpersons -- both English and mathematics -- who culled much more limited lists of proficiencies from the original ones. Teachers would also select and write the items which would comprise the tests to be used on the 1981 graduates. The tests would finally measure no more than 60 proficiencies.

Having experienced continual budget cutbacks due to severe declining enrollment and penalizing state legislation as a result of Proposition 13, Newport-Mesa turned its limited budget and its CBE system to the proficiencies and the support needed to guarantee student success. The district's new Development Lab director, Nola Rochelle, began development of several types of teacher support materials justified by student needs and demanded by teacher requests. Among the instructional materials created were:

(1) Basic Skills Review Packets -- individual review packets for each proficiency containing a pretest, review instruction and exercises, posttest, and (removable) answer key. All packets were directly linked to the item specifications handed out to teachers and distributed widely in the schools.

(2) Proficiency Assistance Catalogue -- a brochure listing both district-written and publisher materials found by teachers to be helpful in teaching to the tested skills. The Development Lab assists teachers in teh procurement and duplication of materials teachers request by phone or district mail.

In addition to assistance to the middle and high schools in the graduation proficiency areas, the Development Lab began work with the elementary schools to articulate the proficiency skills at the lower grade levels -- K-6. In the years 1980 and 1981, proficiency tests were devised for pre/posttest use at all elementary grade levels and by 1982 were mandated by Dr. Loats, the Deputy Superintendent, as a non-optional testing program. Although some instructional materials are available to assist upper grade teachers, very little exists for primary. As it is affordable, more development may occur in this area.

The most impressive CBE support at the secondary level proficiencies remains the creation of a summer school basic skills "lab." The six-week summer program was designed to take advantage of the pretest data available on students in grades 9-11 and to provide a wide variety of instructional materials (including those mentioned previously) and individual student attention for attending pupils. Staff selected to teach in the program attend important inservice sessions where the instructional materials are previewed and instructional approaches discussed. A testing room is set up on the summer school site where students take and retake mini-tests at the discretion of their lab instructors. Although enrichment is provided for students who wish to remain after they have mastered deficiencies, lab instructors have found that students prefer to "take care of business" and leave without the credits available for remaining the full six weeks. Seat time has been so deemphasized with this program that students have little motivation when proficiency requirements have been satisfied.

Beyond the focus on the proficiencies in the basic skills and supporting materials programs, the Board of Education continues to desire the ultimate fruition of the original CBE plans. This year the district will administer the first proficiency test that goes beyond the "minimum" -- both in level of skills and as a district-imposed (rather than State-imposed) graduation requirement. The class of 1988 will be responsible for social studies graduation proficiencies. Like basic skill areas, social studies will become the beneficiary of not only our past proficiency experience but our seventeen-year goals. Through three superintendents and several turnovers in research and development personnel, the blueprint and the vision have remained the same.

Summary

The CBE system of the Newport-Mesa Unified School District has been able to maintain momentum through several changes in school district managing personnel. The Board of Education that exists in 1982 has only one member who was on the Board in 1965. Since unification, there have been three superintendents although the Deputy Superintendent (Dr. Loats) has remained the same during the entire period. Due to the involvement of the staff from the inception of th CBE idea, there remains a widespread commitment to continue the maintenance, revision, and new development needed to keep the CBE system viable. As a result of declining enrollments and budgets, recently-hired staff have been laid off and no new teacher contracts have been tendered since 1975. The majority of the remaining staff is committed to Newport-Mesa's CBE system.

The Newport-Mesa Unified School District has in place a competency-based education program anchored by adopted Board of Education policies. Although the major emphasis of the program at present is in the monitoring of progress made by students on a minimal level of desired proficiencies in the basic skills, efforts are beginning to move the CBE system more into higher levels of expectancy. The program's success has been indicated by several factors:

(1) Teachers have become more involved in individualizing instruction in the basic skill areas.

(2) The yearly performance of students has also improved, both in terms of higher percentages of students demonstrating proficiency on the minimally required skills and also in terms of district scores on norm-referenced tests.

(3) Parents in the district have expressed greater satisfaction over communication from the district about the basic expectations for their children and how their children are doing.

In the process of implementing its competency-based education system, we have learned, sometimes the hard way, a number of important facts regarding the pioneering of such a program in the school district. The following is a brief summary of those facts:

(1) It is possible to define the core of an instructional program in terms of measurable objectives although one must be ready to devote a great amount of resources and energy to do this.

(2) Although more resources, including time, will be expended, it is worth the time to develop a system from within a school district rather than to attempt to adopt a system of instructional outcomes taken from some other district. Staff is much more committed to the implementation of a system if they have had the opportunity to work on its development. "Home grown" products do have their merits. Even though the resources and motivation may not be available to a school district to spell out all of the curriculum in terms of measurable objectives, at least the basic core can be done. Such an understanding cannot be accomplished in a brief period such as one or two years if staff is to be heavily involved.

(3) In the implementation of a competency-based education system, one cannot be too aware of staff concern for how the student assessment data will be used. The goals of such a system must be to assist students to learn. Once these goals are clear, it may be possible to use some of the information to develop evaluation agreements between teachers and principals. An exciting moment in Newport-Mesa's history was when a teacher complained to district staff that she did not see why she couldn't use test results from the SPM program in her Stull Bill agreement (teacher evaluation agreement).

(4) In developing measurable objectives for any instructional program, you need to move into the assessment phase before you can be assured of everyone's attention on the instructional objectives. Once assessment information has been collected, you must be prepared to go back and reanalyze the content and scope of the various objectives that have been developed.

(5) The task of scoring objective-based tests that require a "subscore" for each objective is an enormous one. The use of a computer is almost mandatory.

(6) When implementing a competency-based education system where mastery information is reported by individual objective rather than by total percent of correct items, it takes a very long time before staff truly comprehends and becomes comfortable with this way of reporting.

Questions We're Commonly Asked About Our System

How much did it cost? Approximately $100,000 per year was devoted to developing Newport-Mesa's CBE system for the ten-year period, 1966-1976. From 1976 on, all development and operation costs of the system have been met through normal operating budgets of the various district offices involved.

Was the district's system developed primarily by outside consultants? No. Consultants were used to assist district staff in developing plans for the overall system, but actual work on the system was done by district staff.

Did you buy objectives and/or test items? No. All objectives and items were developed by staff. Consultants were used on occasion to provide inservice training for staff relative to skills needed to develop good objectives and items.

Does your CBE system deal only with "minimum" expectations of the Board of Education, staff and parents? No. Although the majority of required CBE tests in the district are oriented toward "quality control" or seeing to it that all students have mastered required basal skills, some tests are used that measure skills that are not required of everyone but are "expected" of most students.

Pitfalls/Mistakes to be Avoided

In the rush to implement various testing programs, sometimes computer programs were developed in such haste that they did not follow standards normally followed by the Data Center and were not always properly documented. This fact made it more difficult to debug the programs and/or add to them later on.

Although the school district emphasized staff involvement in the development of its CBE system, there were times when feedback was needed from schools that the principal was allowed to "vote" fo his/her staff rather than getting responses from each teacher. This always turned out to be a mistake since more often than not, it was those very same teachers who objected to decisions made from the surveys in which they had had no input.

Factors Contributing to Success

The continued use of district staff in the development of the system proved to be a major asset. Not only did this contribute to the final acceptance of the system by the staff but it also provided the district with a cadre of trained professionals to continue the development work.

Newport-Mesa's CBE system enjoyed, from the beginning, the support of the Board of Education and the Superintendent. The basics of the system are anchored in Board policy and hence will not "go away" overnight.

When turnover in the district staff responsible for implementing and maintaining the CBE system was necessary, replacements were selected who were trained in criterion-referenced assessment and who were familiar with the district's philosophy and policy in this area.

Low turnover of principals and teachers in the district was an aid in that inservice training could be directed toward building on previously-trained skills rather than on teaching the basics.

CLARK COUNTY'S EXPERIENCE

by Theron Swainston

District's Salient Features

The Clark County School District was created in 1956 as part of a legislated reorganization to consolidate all of the school districts in the entire state of Nevada into seventeen districts -- one for each county. Since its creation, the district grew from 20,000 students in 1956 to approximately 90,000 in 1982. The minority enrollment typically comprises 25 percent of the total. In 1979 the State Legislature enacted an extensive property tax cut and placed stringent restrictions on public agency spending and revenues. In 1981 further legislation dramatically shifted the revenue base from real property taxes to retail sales taxes. The major national recession which began in 1981 vividly demonstrated that this "tax shift" from real property to sales tax placed Nevada school finance at the mercy of economic ups and downs for the first time. In 1982 a $70 million school construction bond was defeated marking the first time the District had lost a bond election.

Tax revolts, tax shifts, a defeated bond referendum and a major recession all served to strain the relationship between the community and the schools and confirm the value of an instructional management system that allows teachers, administrators, and board members to demonstrate accountability for acceptable student achievement.

Genesis of a Data-Based Instructional Improvement Movement

In 1969 and 1970 a few central office administrators became increasingly uncomfortable with test scores and instructional inconsistencies. Year after year the norm-referenced test results were below the mean and, in some grades, certain major subtests yielded percentile scores in the twenties and thirties. Initially, low scores were rationalized by statements such as "Those tests don't measure what we teach," or "The norms don't fit our district," or "You can't accurately measure school effectiveness where student transiency is so high." There were also concerns about gross inconsistencies in curriculum and measured achievement from school to school.

At that time, concern for student achievement was kept from the forefront through critical writers and legislation mandating minimum proficiency and demonstrated accountability in neighboring states. Also, at that time, a district administrative reorganization aimed at centralization placed the total supervision of all schools under two assistant superintendents--one for elementary and one for secondary. Thus with the need demonstrated and the organizational structure put in place, the stage was set for a major thrust toward data-based instructional improvement.

Shape and Direction of the Movement

The Assistant Superintendent for Elementary Education, encouraged and assisted by certain other administrators, assumed the leadership for the design and implementation of a comprehensive system for instructional management and improvement beginning in 1970. It was determined that in light of the rapidly changing demographic composition of the district, any significant and lasting improvement would have to be centrally developed and directed. While the system's design and programs were not mandated by the superintendent or the board of school trustees, so long as too many feathers were not ruffled or too many financial demands made, they were usually supportive. At one point, as some principals began to feel uncomfortable with the demands for change and accountability, as they visited with the superintendent they asked him for relief. Initially he sympathized with them, but as he was made more fully aware of the system's potential he became a staunch supporter.

After approximately seven years of development and utilization in the district's elementary schools, in 1978 a new superintendent directed that the system be adapted for and implemented to a limited degree in the secondary schools. This required considerably less development time because a model was in place and operating. However, that superintendent resigned in 1981 and the extent of secondary development and implementation became uncertain.

From the beginning, certain assumptions about curriculum, evaluation, staff development, management and school administration, and the role of educational research provided continuity of approach and design. As the movement extended in years and people, these assumptions became increasingly important. Perhaps the three most basic and far-reaching assumptions were:

1. Goals and objectives need to be clearly written and widely communicated.

2. Means must be provided and used to assess the degree to which objectives are attained.
3. All assessment should culminate in program improvement decisions.

It was always assumed that each component of the overall school and instructional management system would be designed to fit into and facilitate one or more of the above three basic assumptions. Those three came to be expressed commonly as "What should be," "What is," and "How to reduce the difference." It was also assumed that the probability of objectives being attained would be greatly increased if each individual were held accountable through a modified management by objectives and results approach.

Some of the other assumptions that were made consistently were as follows:

1. Curriculum guides should be developed in the district and should define the specifics of curriculum content through learner behavior objectives for each subject.
2. Effective instruction does have research verified characteristics, and instructional programs and teacher performance can be evaluated objectively.
3. Teacher effectiveness can be improved through skilled supervision and relevant staff development.
4. The principal can play a key role as an instructional leader within a school and will effectively fill this role if given the skills, motivation and direction to do so.

Instructional Management System Design

The system has three major components or subsystems. The first component consists of goals, objectives, and evaluative criteria ("What should be"). These range in scope and specificity from a simple math skill in the Kindergarten Curriculum Guide to annual priority goals adopted by the superintendent and the board of school trustees. At the school level the heartof this component is a set of objectives and standards for school operation known as Elements of Quality.

The ten basic elements listed below have become widely accepted and are commonly used as criteria in planning for and evaluating instruction. Each of the ten elements (an eleventh was added for student activities in secondary schools) has sub-objectives or specific evaluative criteria. These specific criteria serve as guides for curricular and instructional planning and both self and program evaluation. The first five also serve as a set of observable criteria to guide principals in evaluating teacher performance. All of the elements are used by principals in making an annual assessment of their school (note Element 6) and by the supervisors of principals in making an annual evaluation of each principal's performance.

Clark County School District

ELEMENTS OF QUALITY FOR ELEMENTARY SCHOOLS

Note: for brevity, the subobjectives and evaluative criteria have either been omitted or printed in abstracted form.

1. Curriculum guides serve as the basis for classroom instruction. Lesson objectives and learning activities are planned in accordance with specified skills and concepts contained in Clark County School District curriculum guides. Content of textbooks, etc., is used selectively to teach and reinforce skills and concepts specified in curriculum guides.

2. Student achievement is commensurate with ability or other established expectancies.

3. Individual differences in the educational needs of students are identified and appropriately met.

The learning needs of students are assessed in relation to established objectives. Students are grouped for instruction according to assessed needs. Instruction is adjusted to student learning rates. Opportunities are provided for students to use their most effective ways of learning. Learning progress is monitored and recorded.

4. Provision is made for the social and emotional development of students. Positive direction and reinforcement are the primary means for motivating students. Students respond positively to teacher model and direction. Students remain attentive to their work. Students display positive behavior when interacting with each other. Students are provided opportunities to initiate, direct, and evaluate some of their own learning activities. Student talk and movement are appropriate for the learning activity.

5. Instructional methods are consistent with established objectives and proven principles of learning.

   The teacher sets a positive climate for learning (stimulates student excitement, anticipation, curiosity, etc.). The teacher clearly communicates lesson objectives and their importance to students. The teaching activities are appropriate for concept, skill or positive behavior lesson objectives. The teacher clearly communicates directions for follow-up activities. Teacher questioning strategies (questions, responses, reaction) facilitate the development of thinking skills (literal, interpretative and critical). Resources are appropriately used for achieving lesson objectives. The physical environment of the classroom is well organized and designed by the teacher to enhance, stimulate and extend learning.

6. A management system providing for needs assessment, priority objectives and plans, monitoring and evaluation by results is effectively used by the principal and teachers.

   Principals are involved and involve their staff members individually in a structured assessment, priority setting, planning, evaluating and reporting process for improving performance results in relation to established criteria.

7. Personnel management procedures prescribed by law, regulation, and contract are effectively administered by the principal.

   The principal makes (and records as appropriate) frequent visits to classrooms to directly observe the instructional program and holds individual conferences with teachers as needed to discuss their priority objectives and plans, their progress toward objectives, needed change and assistance and overall performance results.

8. Staff effectiveness is promoted by the principal through proper application of proven principles of leadership and management.

The principal provides for staff involvement, promotes an open climate, provides training and assistance as needed, reinforces good performance of the staff, and accurately assesses and appropriately responds to measured faculty opinion regarding the operation of the school.

9. Community confidence in the school is established and maintained. Parents are kept well informed regarding the school, its objectives, programs and procedures, are provided convenient means to express their opinion and suggestions regarding the school, are kept well informed regarding their child's school program and progress, and are provided means to be involved at the school and supportive of its objectives and programs.

10. Management organization and procedures for the school are clearly written, effective and consistent with the established procedures and regulations of the District.

Necessary and standard school regulations and management procedures are clearly written in a staff handbook. Provision is made by the principal to monitor and evaluate the management functions of the school to identify exceptions to established standards and procedures and to appropriately deal with management exceptions.

A second component consists of both formative and summative measures of the extent that objectives are achieved ("What is"). Norm-referenced (nationally standardized) tests are administered in grades three, six, eight and eleven. Locally developed criterion referenced tests are used in reading and math in three modes. One CRT form is used as a general placement test generally at the beginning of the school year. Another CRT form is a series of specific diagnostic instruments used to identify specific learning needs for individual students. The third CRT form is an end-of-year measure of mastery of specific instructional objectives. Other measures of "What is" include structured surveys of parents, student, teacher and principal opinions and recorded direct observations of instructional personnel and student behavior utilizing specific observable criteria growing out of the Elements of Quality. The opinion surveys are machine-scored and by computer processing school personnel are provided data through printouts keyed to the Elements of Quality.

The third major component of the system directly addresses the challenge of data-based instructional improvement ("Reducing the difference" between what should be and what is). This component includes a framework for goal setting, priority planning, monitoring and staff development. Utilizing the data base provided by comparing "What is" with "What should be", each level of the district organization from teacher to superintendent and school board selects as priorities for change a limited number of objectives which are realistically attainable and offer the greatest probability for instructional improvement. Each of these priorities is written in a universal format to promote simplicity and common understanding. The format includes statements of identified need, objectives, activities to achieve the objectives, the means for evaluation, the person(s) responsible and finally a statement of evaluation added at the end of the annual cycle. To encourage the selection of significant priorities, a direct relationship between achieving personal priority objectives and personnel performance evaluation is carefully avoided.

Results

Standardized test results have clearly shown that the district's efforts to improve measured student achievement were effective. Elementary achievement at both the primary and intermediate grade levels increased approximately 20 percentile points over a period of seven to ten years from

1970. For example, the sixth grade total reading and math percentile scores had increased from 42 and 38 in 1970-71, to 60 and 64 in 1981-82. The second grade reading and math percentile scores had increased from 50 in 1970-71 to 71, and 76 by 1976-77. In 1978 standardized testing was moved from second to third grade where in 1981-82 the total reading percentile was 64 and math was 71.

Since the secondary schools did not begin to utilize a comparable instructional management system until later, direct comparisons between elementary and secondary achievement are not so meaningful. However, it is reasonable to assume that elementary students would carry some of their higher achievement into the secondary grades. The eighth grade reading and math percentile scores increased from 40 and 36 in 1975-76, to 59 and 65 in 1981-82 and the eleventh grade increased to 53 and 57 by 1981-82.

The results of district-developed criterion referenced measures of elementary student proficiency have been used to group students within a class for instruction, to diagnose individual student learning needs and to evaluate instructional programs in terms of student mastery of instructional objectives. Care has been taken to avoid using the criterion-referenced placement and diagnostic tests as measures of student performance for accountability purposes. By machine-scoring placement and end-of-year criterion referenced tests we can give printouts to individual teachers and principals. These printouts provide extensive analysis data by student, classroom (teacher), grade, school, and district.

An analysis of the responses to elementary teacher and parent opinion surveys which are criterion referenced to the Elements of Quality reveals certain attitude patterns that have been consistently positive for several years. Teachers and parents respond to each item in the survey instrument on a five-point Likert-type scale. The weighted point scores could range from a low of 1.00 (strongly disagree) to a high of 5.00 (strongly agree). From 1978-79 to 1981-82 the mean weighted response on the teacher instrument increased from 3.36 to 3.46 for Elements 6 through 10. The mean weighted response on parent survey remained very stable, but high with a mean of 4.36 in 1978-79 and 4.37 in 1981-82. Certain items on each survey have special significance. For example, on the parent survey, "The teachers use effective methods and materials to help my child learn" yielded 4.46 in 1978-79 and 4.47 in 1981-82 with an average 14,000 parents responding, representing all of the 71 elementary schools. Two of the most significant items on the teacher survey were "My teaching performance is accurately assessed relative to District-estabished criteria" (3.51 in 1978-79 and 3.55 in 1981-82) and "The principal's supervision of my classroom performance results in improved instruction" (3.29 in 1978-79 and 3.44 in 1981-82). An average of 1,950 teachers responded which was nearly 100 percent of the total number of elementary teachers on staff. One of the challenges of a highly structured and centralized instructional management and accountability system such as ours is managing the affective response of teachers. The teacher opinion survey is used annually to allow individual principals and district administrators to monitor the attitudes and feelings of teachers. Both the teacher and parent surveys are administered in a manner to encourage objectivity through anonymity of respondents. Since each item in the survey instruments (35 for teachers and 15 for parents) is a clear statement of a desirable condition or objective to be achieved, they represent a directing force as well as a means to measure attitudes.

A Unique Feature

A very significant and somewhat unique characteristic of the Clark County experience deserves special consideration. Many top level Clark County administrators believe that it is the single most vital difference between what might have been a passing experiment of limited impact and one that has endured and has had a profound influence in dramatically reshaping the instructional program of a very large school district. The management of all instructional improvement efforts was deliberately linked with the overall direct management of schools. The following examples will serve to illustrate this point.

When one independent department was responsible for developing curriculum and others were responsible for implementation, the use of curriculum guides was considered to be optional and many principals and teachers did not take them seriously. However, when key line administrators assumed responsibility for the development, implementation and success of the curriculum, a new standard was set. One of the criteria used to evaluate the performance of principals and teachers was the extent that curriculum guides were used as a basis for instruction. (See Element of Quality No. 1.) A uniform curriculum became a reality.

Prior to the early 1970's, test scores were not taken seriously except by a few people and mostly by the central office staff responsible for testing and evaluation. However, when the comprehensive school and instructional management system was developed, acceptable test scores became one of the criteria not only for evaluating instructional programs, but also for evaluating the performance of instructional personnel. (See Element of Quality No. 2.) Here an important distinction was attempted. An adage was coined that "When tests are used as weapons of supervision rather than tools of instructional improvement, they soon become useless for either purpose." Rather than an individual's effectiveness being evaluated upon reaching a norm or an absolute standard, they were evaluated upon showing needed or significant improvement. This permitted teachers and principals in traditionally low achieving areas to be winners and it required those who happened to be in traditionally high achieving areas to do more than exceed the district means. It was reasoned that if the bottom line of the school business is student achievement, those who control the day-to-day operation of the business should be directly accountable for the bottom line.

It was determined that principals and their direct supervisors would be the first line instructional leaders. They would plan and conduct or direct staff development for themselves and teachers. They would be involved in curriculum development and all curriculum implementation and instructional improvement programs would flow through them.

Challenges and Recommendations

In a continuing sequence of development, implementation and refinement since 1969, the aspect of the system which currently offers the greatest challenge and promise is staff development. For example, considering the classroom teacher and school principal we can now describe quite well the characteristics of good performance ("What should be") and with considerable accuracy assess the performance of a given individual skill and motivation. An additional challenge of a highly structured system is providing for upgrading and renewal of all the components to ensure that it is dynamic and relevent.

Certain recommendations seem to be appropriate for others who might consider utilizing any part of the Clark County system or experience. It is important to assess the philosophy or assumptions of key leaders to ensure that there is compatability with goals and approaches that are developed. Objectives are most basic and should be clearly written for all levels of the organization and generally accepted by those responsible for attainment. There are clear time advantages of adapting existing models where they can be found rather than starting with nothing. Formal endorsement and support from the Superintendent and the Board of Trustees is ultimately necessary and should be sought early. Heavy involvement in design and development by those staff most directly affected helps to ensure acceptance and commitment.

Questions Commonly Asked About the System

Administrators who work with the Clark County system are most often asked if the management and accountability system has made a positive difference in teacher and principal behavior and student achievement. Principals and teachers are much more structured and objective in planning and evaluating instruction as a direct consequence of the system. While it is not always possible to quantify the exact degree, it is possible to demonstrate that the system is the most significant cause for dramatic improvement in test scores. our principals and teachers are now far more more sensitive to the need to ask questions of themselves such as "What changes can we make which will result in the greatest possible improvement in program quality, and how will we demonstrate or measure the desired improvement?" This represents not only a more objective attitude and approach pattern, but has led to a dramatic increase in planning and evaluation skills and knowledge at the school level.

The question of whether the increased centralized structure has reduced individuality is often asked. Great care has been taken to place emphasis upon specifying curricular objectives and measuring results rather than limiting methods and individual techniques. For example, the district carefully avoids single textbook adoptions and allows teachers and principals in individual schools to select texts from an approved list of two to five books for each subject. Even wider individual discretion is allowed in the selection of supplementary instructional materials.

The question of how widely the system is accepted is also raised. The fact that the system has spread from elementary to secondary schools and has endured at the elementary level for more than ten years says something about its acceptability. However, the system, has been and continues to be resisted by certain people. The line of people demanding or volunteering to be held more accountable is never long. Approaches as such annual opinion surveys and widespread involvement in the development of system components have increased the acceptability of the system. The system survived a lawsuit by an employee group which originally saw it as a threat to employment security.

Pitfalls/Mistakes to be Avoided

Perhaps the most serious pitfall to be avoided is the tendency to use such a management system almost entirely as a tool (or worse still, a weapon) for controlling at the expense of planning, organizing and motivating functions. To avoid this, great care must be taken to ensure that those who administer the system understand and apply the principles of management and leadership upon which the system is based and without which the system will die of its own weight.

Another pitfall to be avoided is moving too far too soon. Components of the system should be structured in detail and field-tested before widespread implementation is attempted. A master plan for the entire system is highly desirable in order for each component to be seen in proper perspective. For example, a criterion-referenced test for sixth grade mathematics should not be viewed as boards and nails, but as part of a building that can be viewed in blueprint elevations and floor plans. To move too far too soon sets up resistance shockwaves stemming from misunderstanding and anxiety about the value of such a system to either the employee or the students.

Special attention should be given to keeping paperwork to a minimum and the major focus on what happens in the classroom compared with what should happen.

Factors Contributing to Success

I believe that a major factor contributing to the success of the Clark County system is consistency of design and application. Because the system rests upon a clearly defined philosophy and research-proven principles, it has weathered well the shifting winds and storms of criticism and passing "latest" innovations. For example, when "Theory Z" and "Quality Circles" came upon the scene, the system had a place for them. When the Beginning Teacher Evaluation Study became a major topic of discussion, the system easily accomodated the major findings and implications.

The overall system design rests upon a core of overriding criteria defining quality at the school level, and these Elements of Quality are manifest throughout the entire system. We measure almost everything is measured in light of its consistency with these criteria. The application and operation of the system rests upon certain proven principles of management, and administrative behavior is evaluated in light of consistency with these principles (process) and the criteria defining quality education (product). Because this consistency has been preserved over an extended period of time, the system and its components have become second nature to those who have "grown up" with our system.

VALLEJO'S EXPERIENCE: LINKING TESTING WITH INSTRUCTIONAL IMPROVEMENT

by Joan McDonald

Introduction

The Vallejo City Unified School District is a medium-sized California school district (14,000-15,000 students) located in an urban, multi-ethnic community. Over 50 percent of the students are members of a minority group. Black students represent the largest minority population, followed by a fast growing Filipino population and an increasing Hispanic population.

Vallejo is a low-income community. Mare Island Naval Shipyard is the major employer in the city; the school district is the second major employer. Many of the students' parents are unemployed or have low-paying jobs. Most students enter school without preschool experience.

Vallejo's students have many needs, both academically and economically. The primary goal of the district is to provide the students with an excellent education and to see that they reach their fullest potential. The same goal probably exists in all school districts. This is the story of one district's efforts to make the vision a reality through instructional improvement. The use of test data became a catalyst for analyzing instructional problems.

The vision that the superintendent and his staff have is a district with a clearly expressed philosophy and direction. Too often a school district does not take the time to articulate a philosophy and to establish and communicate clear directions both within the district and to the public. It is assumed that everybody wants the "best for kids" and knows how to bring that about. Part of Vallejo's vision has been to establish and articulate a clear sense of purpose.

Improved instruction was a part of the vision. The district's definition of improved instruction included: 1) a district curriculum where objectives, learning activities and evaluation procedures match and are consistent; 2) schools in which principals and teachers engaged in ongoing efforts to analyze and improve teaching; and 3) the use of test results to improve instruction.

A long-range plan for curriculum development was initiated. A district professional development center was established to research, teach and promote sound instructional practices. Ongoing efforts were made to become familiar with the purposes, content and uses of tests and test results. This article will describe in chronological order the problems that precipitated the actions taken, a description of some of the solutions and a summary of lessons learned.

The Beginning

The California Assessment Program (CAP) was a major influence propelling the Vallejo School District to use evaluation information to improve instruction. CAP was initiated in 1974, toward the end of an era of decentralization in the Vallejo School District. Between 1965-1975, school sites had considerable autonomy and had assumed responsibility for curriculum development and staff development. Schools had experimented with a variety of organizational schemes such as team teaching, departmentalization and non-graded classrooms and had tried various schedules for students. During this decade, decentralization was actively encouraged to promote school-site autonomy and decision making. District-wide coordination took a back seat to experimentation and the development of school-site ownership of instructional programs.

When CAP came on the scene, it assumed center stage. Test scores made front page headlines and became "the agenda" for disgruntled members of the governing board and the community who were already critical of the schools. CAP was a catalyst. The educational community had no choice but to deal with CAP, though initially it would have preferred not to.

Problems That Arose

Dealing with CAP testing and its results productively was difficult for us for several reasons. First, the previous decentralization of our school system impeded any coordinated and comprehensive effort to use the results. Second, CAP testing became a highly emotional issue both with the public and within the school district. Third, our school staffs had limited knowledge about how to use test results to improve the instructional program.

The previous decentralization of the school district contributed to our problems because there were no consistent district-wide instructional practices and clear curriculum outlines. Additionally, schools had become accustomed to making decisions for themselves without necessarily coordinating their efforts with other schools. No district-wide statement of the curriculum existed to tell teachers what skills should be addressed when and in what formats. Title I project elementary schools had done more in terms of defining the curriculum for themselves, but they had worked in isolation from one another. So, in fact, each Title I school used a different skill continuum. Non-project schools typically had no curriculum frameworks.

Textbooks provided the curriculum, and teachers taught what the teacher's guide prescribed and used the methods described there. A further problem was that schools made autonomous decisions about textbook selection and, in fact, many different series were being used throughout the district.

The emotionalism surrounding CAP interfered with its being used as a tool for improving instruction. Some of the most severe critics of schools seized on CAP as a vehicle for casting criticism and blame on the schools. The educational community responded defensively to past and current misuses of test information. Ignorance increased emotionalism on both sides.

School staffs had limited knowledge about the purpose, content and usefulness of the CAP tests. The district leadership had not emphasized the use of test results. Since the district curriculum was not defined, principals did not have identified expectations to help them use test results to improve the school's instructional program. Neither principals nor teachers had training to equip them to properly analyze, interpret and use test results for program planning.

Central Office Decisions About the CAP Test: School-site Analysis

The central office leaders decided that a district-wide effort to use test information to improve instruction had to be initiated. The plan for accomplishing the goal included developing awareness on the part of principals, training principals in the use of test results, and providing direction for school-site analysis and planning. This process led to a series of long-range efforts in the area of curriculum and instruction.

A centralized plan for using test information was developed to initiate some consistent processes across the district for dealing with test results. We emphasized that those in central office leadership roles believed that test data was important to program development and evaluation. As will be described in the following sections, our plan moved methodically from promoting awareness, to training, to requiring application of knowledge at the school site. Use of test data was the intended outcome, but those who were providing leadership were aware that some emotional and awareness issues had to be addressed first. Although site managers did not respond enthusiastically in the beginning, a growing appreciation for the value of using test data did develop. In the central office, we initiated a process to ensure that school-site staffs used test data to improve the instructional program.

The first step in developing principal awareness about the power and uses of test results was to overcome the emotionalism that surrounded the issue. The most immediate reaction of principals was highly defensive and resistant. Without even reviewing the test results of their schools, a majority of the principals rejected them. Principals claimed loudly, "CAP didn't test what we teach," CAP is "biased" and "irrelevant." Another favorite cry was, "CAP was created by test makers who have little or no knowledge about schools and what we do." The fact that the district had a low percentile ranking even in relationship to other districts in its comparison score band didn't help a bit. The fact that the local newspapers headlined stories that governing board members were accusing the schools of failure only served to further hamper communication.

The district leadership realized that further emotional discussion would only exaggerate the problem. We made plans to provide training to principals, selected teachers and the governing board so they could begin dealing with CAP productively. Inservice topics focused on information about the test itself, skills to analyze test data and strategies for interpreting and using results to plan for instructional improvement.

The central office administrators contacted the Office of Program Evaluation in the State Department of Education. Consultants from the State office came to Vallejo, met with the district staff and principals and conducted a series of awareness-level sessions on the purpose and development of CAP. Test score reading and interpretation were also essential components of the awareness level training provided by the state staff. The information contained in the workshops was essential in terms of providing a knowledge base to begin dealing with CAP results, yet the training itself did little to affect the defensive attitudes of principals and teachers.

The next activity we initiated had the most powerful impact on school staffs. District office administrators drew up a three-step process in which school staff were required to work through and submit in writing an analysis of their test data and a plan of action. The process involved becoming familiar with the content of the CAP test, then analyzing the school's program and finally identifying directions for improvement. The directions were then reported to the governing board.

This school-site analysis process required school staffs to use the Test Content Specifications provided by the State Office of Evaluation to become familiar with the actual skills tested by CAP. The purpose of this component of the process was to have teachers and principals answer the question, "Are we teaching what is being tested?" This was an important step because it caused principals and teachers, for the first time, to deal specifically and objectively (rather than emotionally) with what CAP tested in relation to their school's instructional program. The argument that "this doesn't test what we teach" faded abruptly.

ONE SCHOOL'S STORY

The following is one principal's description of how the school-site analysis process worked at his school. The school described here was actually located on two campuses and served by one principal. The student population was made up of over 70 percent minority students and a high percentage of students who demonstrated low academic achievement on standardized tests. At that time, the school also had the highest percent of AFDC (Aid to Families with Dependent Children) students.

The Principal's Description

Most of the school-site analysis activities were related to the use of CAP results. Other standardized tests were also administered to evaluate our Title I program, but the use of other test data at this point in our history was primarily a result of transferring knowledge gained by using CAP results to other testing programs.

The Instructional Associate (a non-classroom teacher paid for out of Title I funds), selected teachers and I reviewed the test results to identify areas of concern. Areas of concern were defined as those areas that either seemed to be especially discrepant with overall test results or those areas where their results seemed discrepant given the amount of instruction that had been provided and the results expected. It was also decided that no more than three areas would be selected in reading and three in math for a more in-depth analysis.

The Instructional Associate and I used the State manuals and other materials we received during the training provided by the State consultants to:

- identify and define the content being tested in the areas that had been identified;

- restate the learnings that were being tested in our own words and in terminology familiar to our staff;
- prepare a task analysis of each of the skill areas.

The above information was shared with all staff members at regularly scheduled meetings. A decision was made to familiarize all of the teachers with CAP, not just the teachers of students in the grade levels that were tested. We wanted the entire staff to share the responsibility for improving our instructional program.

After the skill areas tested by CAP were carefully analyzed and shared with the staff, they were matched with the appropriate objectives and evaluation instruments in the reading and math curriculum management systems being used at our school. Some of the teachers were involved in validating the match between the skills areas being tested by CAP and the skills being taught through the use of the reading and math textbooks at the school.

All teachers worked individually or in small groups to develop prototype lessons and activities for teaching the identified skills at different levels of difficulty. These products were shared among the teachers for use in classroom instruction. The lessons and activities were not the only results of this part of the process; it also produced a renewed sense of purpose and teamwork among staff members.

In addition to improving their lessons for students, teachers also began working on teaching their students test-taking techniques. Teachers examined the testing formats used by CAP (multiple-choice, fill-in-the-blanks, etc.) and incorporated those formats into the practice activities they built into their lessons. Teachers also familiarized themselves with some of the key words used in test directions and taught them to their students. The Instructional Associate and I developed practice activities and conducted practice sessions for groups of ten to fifteen students.

Outcomes at the School

A very important outcome of the school-site analysis process at the school described here was the change in attitude toward testing. Staff meetings were used to discuss the value of using a testing program to plan instructional improvement. Teachers came to the conclusion that testing is an important component of a school program, that analyzing test results could help them monitor and modify instruction and that test taking itself is a valuable skill for students to acquire.

Test scores at the school described did begin to improve as a result of the school-site analysis process. Miracles did not occur. Test scores were not immediately and dramatically much higher. The instructional program did improve, the school staff developed a more positive attitude toward testing and a more cohesive plan for using test results, and test scores reflected those efforts. The principal's active leadership, interest and involvement provided a powerful message to the staff and led to their active involvement in using test data to improve instruction.

District Outcomes Leading to New Directions: Curriculum, Staff Development and CRT's

There were several outcomes that resulted from developing a greater awareness about CAP, providing training to help principals and teachers use test results for instructional improvement and requiring each school to use a school-site analysis process.

The first outcome was an improved attitude on the part of the educational community with regard to the potential uses of test information. The educational community as a whole became more objective, analytical and confident as it developed an understanding of information and the skills to act on this information.

A second outcome was a growing awareness and concern for continuity, consistency and coordination of the curriculum that surfaced both within schools and across schools. School staffs also became aware of specific skill areas that needed to be given more attention because students were demonstrating difficulty with them on tests.

The outcomes described above led to some new directions for the district. The realization that there was no well-defined district curriculum resulted in a substantial curriculum development effort.

Within three years district continua in reading/math/language were developed. The district also established a Professional Development Center to provide extensive training to principals and teachers in improving their instructional skills. Attention was given to improving testing procedures throughout the district. Two years later criterion-referenced tests and student profile cards were introduced to complete the curriculum management system in those skill areas.

Curriculum Development

The closer look at the instructional program that followed, a greater emphasis on using test results made everyone very aware that no district-wide statement of what skills were to be taught to whom and when existed. With the advent of the "Hart Bill" proficiency test movement, the district moved to develop the necessary district-wide statement of skills. Skills continua were written in reading, language and math, including the identification of proficiency skills developed throughout the K-12 curriculum in basic skills.

For the first time, teachers came together across grade levels and hammered out the continua. Certainly, those who were most centrally involved in that effort learned the greatest amount from the experience. However, as those continua were introduced to staffs and field tested, a wider and wider segment of the teachers was affected. The first steps toward improving instruction had been taken, as throughout the district, individual teachers were becoming more specific and articulate about the skills being taught. The greatest benefit was a beginning recognition of a move toward a district-wide statement of curriculum.

The next step was the development of the Criterion Referenced Tests (CRT's) in relation to designated essential skills in the continua. The requirement that teachers test their children and record their progress on a district-wide (K-6) student profile card finally moved the continua into prominence as the basis for instruction. The use of the CRT's and requiring teachers to record skill mastery on student profile cards can be identified as crucial to the institutionalizing of the continua of skills. Prior to that time, only some of the teachers truly felt bound to teach the district curriculum.

Inservice training for teachers was provided on school sites by central office personnel assisted by teachers who had been involved in the development and subsequent revision of the continua and the CRT's. The more involved a staff became, the more aware they became of the need to review together the task analysis and instructional strategies involved in teaching the essential skills. Now, staffs are identifying their instructional areas of need through CRT's as well as CAP and CTBS. In addition, they are also identifying areas of need for staff development. Instructional improvement is beginning to show results in test scores.

Principals became aware of the complexity of truly bringing about change as they provided leadership to their staffs in implementing the district curriculum management system. One principal tells the following story which illustrates some of the challenges that confronted her and the solutions she tried at her 4-6 school.

ONE SCHOOL'S STORY

The Principal's Description

> In the spring of 1981, the district curriculum management system package arrived complete with continua, CRT's and student profile cards. Our staff spent the spring and summer playing with a variety of ways to introduce and teach its use to our staff. At this point, we were primarily dealing with the mechanics of using it. How would teachers get CRT's? How could teachers manage 90 profile cards? How could we insure that teaching would go on prior to testing? How could we streamline and facilitate the duplication and storing of CRT's?
>
> We finally developed what we considered to be an excellent, almost foolproof system for management, and an enlightened inservice program with audiovisuals, songs and dances, etc., on how this foolproof system would work was developed and presented in the fall. The inservice was heralded as a great success. Teachers seemed delighted with the organization and toted their boxes of 90 cards (one for reading, math and language for each student) to their rooms filled with enthusiasm and good will.
>
> The first hint of any disaffection was evident in November of 1981 when teachers were required to turn in their student profile cards for review by the principal for the first time. Comments like, "If I could use my time teaching instead of marking cards, my kids would learn something" were common. My review of the cards showed haphazard marking, no uniformity of skills addressed, and general lack of care taken in their marking. This response came from a generally cooperative, hard working, concerned and dedicated staff.
>
> I began a series of informal chats with a variety of teachers during which I asked things like, "How do you decide what skills to teach?" "What does introduced mean to you?" "Do you think you can cover all skills required for your level?" "Which skills are most important to your grade/class?"
>
> By listening carefully and reading between the lines a bit, it became increasingly clear that my group of dedicated upper grade teachers were having a terrible time. The complaints and concerns boiled down to the fact that for many of them it was the first time they were being

forced to teach reading, language and math skills and they weren't any too sure what these skills were and how to teach them. So long as they followed the textbook, they were fine. All of a sudden they had skills to teach and no reference to a specific page number in a book. For many, particularly in reading, it was the first time they were confronted with subdivided content skills. The component parts of reading were all of a sudden isolated for them and many had never taught these component parts.

We realised that in order to get back on track and make the management system an effective tool for us, we needed to do several things. We needed to:

- narrow the laundry list of skills contained on the continua so that teachers would learn to teach skills in digestible hunks;
- select skills for this narrowed list that were critical to mastery of each subject;
- select skills that were consistently critical across the three grades we serve:
- select skills that were measured on most of our assessment instruments;
- help teachers value these skills;
- help teachers learn how to teach these skills.

The steps we took to accomplish these tasks were to:

- use standard reading, language and math committees to discuss and isolate 10-15 skills that were critical and that were necessary at each grade level;
- select reading as our focus for spring 1982 with language and math to follow in the 1982-83 school year;
- cross-match skills selected by the reading committee to the test specifications for Hart Bill, CTBS, and CAP;
- select the 10 skills most commonly measured and critical to mastery of reading;
- divide the staff into three committees with each being responsible for:
  - refining the district's task analyses for each of 3-4 skills;

- developing additional CRT's for grade levels where none were available for each skill;
- developing/ordering/adopting teacher support material for each skill;
- sharing task analysis content, CRT's and materials for each skill with the rest of the staff;
- identifying staff resource persons for each skill.

We were able to refine the task analyses, develop additional CRT's and order materials by the end of the year. Sharing the task analyses and identifying resource persons are agenda items for September 1982 faculty meetings. We will then begin the process again for language in October of 1982. Teachers will be required to begin to complete profile cards for reading in September. As we complete language and then math in the spring, they will begin to complete cards in these areas.

One of the terrific things that happened in this process is that by narrowing our focus to critical, often measured skills, teachers were able to focus on only a few skills to learn how to teach at once. They own the program now because they have developed it and taught it to each other. Their confidence is high and they've painlessly learned new skills. Next steps, of course, include expansion of skills to be addressed and extension of resources and teacher skills.

One critical point was that by cross-matching the skills we selected with the instruments we use to measure skills (CTBS, CAP, Hart Bill), the skills were validated for teachers and I can be sure our students are being taught the skills on which they will be tested.

Staff Development

Curriculum development and staff development go hand in hand. Curriculum development is staff development. The teachers, principals and central office staff members who designed, developed and are implementing the curriculum management system were definitely engaged in professional development. The activities described in this section are specifically the staff development activities that occurred at the district Professional Development Center.

The Vallejo Professional Development Center was established in 1976 with state funds and by law served only the 10 Title I elementary schools. In 1978 federal Teacher Center funds allowed the Center to serve all of the district's schools. The district staff development program was used to help train teachers and principals to make effective use of test data and to improve instruction.

The Professional Development Center program identified specific skills areas to feature in its task analysis/lesson design training sessions. "Identifying the main idea," "inference," "differentiating fact from opinion," were among the skills areas the district identified for focus. The task was essentially to improve the job being done in district classrooms in teaching these skills.

Once teachers received training, they received follow-up visits in their classrooms. Principals, trainers and other teachers observed the teacher putting new skills into practice in the classroom and provided feedback. The follow-up component of the training assured that new skills would be applied and promoted ongoing teamwork and dialogue about effective instruction.

Improved Testing Procedures

In addition to curriculum development and staff development efforts, procedures were instituted to improve actual testing procedures. A coordinator was appointed to organize the testing program and to work with school sites to implement it. In addition to CAP and CTBS testing, a major task was to develop, pilot and adopt proficiency tests. Because of all the work that preceded proficiency testing movement, the district was prepared to address the new challenge in a systematic way.

The development of the district continua was described earlier. Subsequent to continua development, three teachers were released from their secondary teaching assignments to help other teachers incorporate the skills on the continua and tested by the proficiency test into their courses. The use of experienced classroom teachers to help other classroom teachers proved effective.

Principals' meetings were also used to improve testing procedures. At key points during the year, the importance of testing was reviewed with principals. They were asked to come to meetings prepared to share strategies that worked for them to help students and teachers take testing seriously, to schedule testing at times when there would be a minimum number of interruptions and to use test results in a planful way. Principals learned many valuable strategies from one another. The superintendent chaired the meeting as he does most principals' meetings. His clear direction that effective testing procedures were important re-emphasized a district priority.

Where We Are Now

The directions set as a result of the initial concern about testing continue. Curriculum development, staff development and improved testing procedures remain priorities. The past six years marked a period of very active development. The current thrust is to thoroughly implement and refine what exists.

A great deal of energy went into the developmental phase of the curriculum products. The current thrust is to see that the products are used rather than shelved. In an earlier vignette, one principal described the difficulties of actually using CRT's and profile cards. It is a slow, painful process. Each principal submits annually an objective to the superintendent for providing the leadership for his or her staff to proceed to the next level of use of the system. The Curriculum Director provides direct assistance to principals who are evaluated by the superintendent on the degree to which they accomplish their objectives.

The staff development program will include more and more training sessions on how to apply effective instructional practices to using the district's curriculum products. The Staff Development Department and the Curriculum Development Department are working together to find teachers who have worked out solutions to typical curriculum development problems who can train other teachers. The follow-up component described earlier continues to emphasize and assist in the application of new skills in classroom instruction.

Testing procedures continue to be improved both through the central coordination of efforts by the Coordinator of Research and Assessment and through the site leadership of principals. Principals' meetings continue to be used to help train principals and maintain a focus on the effective use of test results for instructional improvement. Without involving school staffs, a curriculum may only be words printed on pages. In order for a curriculum to be taught, it must be wrestled with. Teamwork is an essential ingredient of success and teamwork, too, takes time. A real success is the current level of teamwork.

Planning is another key ingredient of success. An extensive effort to revalidate district goals with school staffs, students, parents, and the community was recently developed. Without a long-range plan and a timeline, it is too easy to drift away from directions. The plan was developed with input from many teachers and principals. It will be presented to all school staffs, parents and the public. An Instructional Improvement Council consisting of administrators, teachers, classified employees, parents, students, and community members will monitor the plan.

A final success has to do with morale. Teachers and principals are probably working harder than ever before. They are tackling serious issues in a serious way. The teamwork, sense of purpose, and sense of accomplishment has had a positive effect on morale. The danger always remains that dedicated professionals will work too hard, too fast, and too long and burn themselves out. The purpose of the five-year plan is to maintain the vision and intensity but pace the efforts in a simple and satisfying way.

Lessons Learned

Many lessons have been learned in Vallejo. Experience has taught that the most important endeavors require a great deal of time and commitment. Our experience has shown that there are some possible pitfalls, some questions that are frequently asked, and some successes to share.

Possible Pitfalls

There are pitfalls to avoid. Our experience would lead us to caution other districts not to attempt a comprehensive improvement effort unless they are willing to allocate the necessary time and human and finanical resources.

The greatest cost is staff time. In Vallejo, the Administrator for Instructional Services, the Curriculum Development Director and the Professional Development Director have allocated a substantial amount of their time to the efforts described here. Release time has also been a substantial expense both for curriculum development and training activities. The time period in which Vallejo's efforts were initiated was one in which state and federal grants were more readily available. The district's commitment was essential but some aggressive grant writing efforts also yielded some financial resources that probably allowed the efforts to develop more fully in a shorter period of time than would have been possible with no external sources of funds. A real pitfall would be to undertake a serious change effort without examining the costs.

Questions Others Ask

We are often asked about the balance between leadership and involvement. What we have found is that leadership and structure is appreciated as long as clear and relevant opportunities for involvement are built into the process. The superintendent and his staff are very much in charge in terms of establishing directions, but school-site staffs have the flexibility to tailor district directions and programs to their needs. Both the curriculum development and staff development efforts have built in a wide variety of ways to be involved as well as varying degrees of involvement. Some teachers and principals are involved only in school-site efforts while others sit on the Instructional Improvement Council. Some teachers and principals choose mainly to receive and apply training in their own job situation while others stretch themselves to receive and apply training and then train others. There are many avenues for individuals and groups to participate in the instructional improvement efforts.

Factors Contributing to Success

The importance of district level leadership is constantly revalidated. The superintendent sets clear priorities, evaluates principals on their ability to carry out those priorities, and directs his staff members to use their resources to assist the school sites in their efforts to address the priorities. The superintendent's vision is an important prerequisite.

A second factor contributing to our success has been the willingness to take the time needed to accomplish change. A curriculum management system takes a long time to develop and even longer to implement. Involvement takes time, but it is time worth spending.

THE SAN JUAN EXPERIENCE: CREATING AN EFFECTIVE EVALUATION/PLANNING MODEL

by Larry Crabbe

Introduction

The areas of program evaluation and planning are relatively new to education. Much of the attention given to them in recent years has come as a result of legislated requirements attached to such specially funded programs as California's Early Childhood Education (ECE) and School Improvement Programs (SIP) as well as the indomitable ESEA Title I Program (now ECIA, Chapter I).

Ten years ago, the San Juan Unified School District created the position of Evaluation Specialist specifically to carry out these mandated evaluation and planning responsibilities. This was the beginning of an exciting experiment which has produced dramatic changes (improvements, we hope!) in the ways in which our schools' manage their educational affairs.

Our evaluation/planning model is one which takes place "where the action is," at the school-site level. It also involves those with high stakes in its outcome, namely, the school's staff, students and parents. Significantly, our model is one which begins with a broad data base. Most importantly, it goes on serving throughout the school year, long after the formal process of planning has been concluded.

In the early seventies, our model was used in fewer than a dozen district elementary schools. As the number of programs expanded, so did the application of our model. It is now in effect in over 40 of our district's elementary, intermediate and secondary schools. We are proud that all of our evaluation and planning support activities, and, in fact, many additional activites, have been provided to all these schools with no addition in staff (a teacher and myself) from the days when less than a dozen were involved. In fact, our entire Research and Evaluation Department has, throughout its history, remained exceedingly small with a specialist staff of four and a budget representing a fraction of one percent of the district's overall budget.

The most important outcome from the evaluation/planning process has been instructional programs which are tailored to the actual needs of the students. These programs are "owned" by teachers and, therefore, are supported by those charged with their implementation. The most important by-products at the school-site level have been a strong sense of school-community relationship, as well as feelings of openness and self-confidence.

In the succeeding pages I will describe our model. I will indicate its philosophical basis, its major components and its chronological evolution. It will be helpful, I think, to "set the stage" by discussing the historic and current contexts in which our evaluation/planning model has spent its youth.

Historic and Current Context

When James Marshall discovered gold at Sutter's Mill in 1848, Sacramento, California consisted only of Sutter's small fortress. That was all. The infamous gold rush that followed instantly transformed Sacramento into a thriving city of considerable size. Many of the other towns to which that great rush gave birth slowly faded with the demise of the rush. However, Sacramento, because of its strategic location at the juncture of the two largest rivers of the great central valley, continued to prosper and grow. This trend has continued, nearly uninterrupted, to the present day.

As Sacramento grew, small communities began to dot the surrounding countryside. As the years passed, they formed their own school districts to attend to the educational needs of their children. Six such school districts, twenty-two years ago, combined to form the San Juan Unified School District. With the passage of time, these individual communities and their district offices merged to form the north-eastern suburban area of Sacramento.

San Juan is now the seventh largest school district in California. It serves over 44,000 kindergarten through twelfth grade students and spreads its influence over an area of approximately 75 square miles. The district as a whole has been described as "middle" to "upper-middle" class. Socioeconomically, San Juan exceeds 80 to 85 percent of California's unified school districts. Sacramento, the capital city of California, owes much of its economic existence to federal, state and local government as well as to the military and aerospace. These "industries" utilize the services of many professional and paraprofessional employees. After work, many of these people come home to the vast "bedroom" community served by the school district.

As might be expected, the district's students have strong academic orientations. Student performance, according to the state's testing program, is in the 73rd to 87th percentile range. Three out of every four graduates attend college. There has been a gradual increase in racial and socioeconomic variety. The current racial composition is 1.9 percent

American Indian, 2.5 percent Asian, 1.9 percent Black, 4.6 percent Hispanic and 89.1 percent Caucasian. The percentage of students receiving public assistance (AFDC) varies widely from school to school, from a low of less than one percent to a high in excess of 40 percent.

In recent years, the district has faced several significant challenges. Our most enduring problem has been financial. Because of our suburban character, the district's tax base and, hence, resulting revenues have been and continue to be low. The community's support of public education has resulted in a relatively high tax effort. But despite this, the district's expenditures still average $30 to $150 less per pupil than the median district in the state. This situation is now complicated by the state of the economy and the discontinuation of funds for special programs. District costs are escalating while, at the same time, funds are drying up.

A second recent challenge to the district revolved around its board of education. Our traditionaly peaceful district has in the past several years received much statewide attention as a result of, not one but two board recall elections.

The third challenge is declining and shifting enrollment. Like many in the state, we are experiencing a gradual decline in enrollment. Were this not enough, our district situation is compounded by a shift of students from west to east within our long and narrow district. As a consequence, schools in the older western portion of the district have surplus capacity while those at the opposite, eastern end are bursting their seams. Solutions to such problems, especially in districts of this size and wealth, do not come easily.

This, then, is the context in which we labored to develop our approach to educational program evaluation and planning.

Our Philosophy

I mentioned earlier that the district responded to early legal evaluation and planning mandates by creating a position, Evaluation Specialist, with those mandated activities as its job description. This decision, possibly more than any other, may have been responsible for the success that the district has had in the area. This may have less to do with the individual filling the position, me, than it does with the fact that a position was created with evaluation and planning as its direct responsibilities. We all, quite reasonably, concentrate the majority of our efforts in those areas upon which our continued employment most directly depends. We also, quite reasonably, limit the amount of time we devote to peripheral duties so as to ensure that our primary responsibilities are not neglected.

Purpose of Evaluation: Planning

Throughout the evolution of our evaluation/planning model, we have held true to several fundamental beliefs as to the reasons for evaluation and planning. We believed that the evaluation and planning model would deliver the greatest "payoff" to the district's schools if these were its primary goals:

- to improve the quality of the local educational program;
- to encourage the most effective or efficient use of scarce financial resources;
- to assist with the attainment of the goals of the school, the school district and the special program(s) mandating the evaluation/planning activities.

There was another, secondary, goal for evaluation and planning:

- to create a written document or "plan" in which to preserve the most significant findings or conclusions emanating from the evaluation/planning process.

If well attended to, this secondary concern can assist schools significantly to move toward attainment of our primary goals, even though it should never, never, exist as a goal or end in itself. A well written plan or "grant" can certainly have a favorable financial impact on the school or district. But, an impressively written grant does not ensure that the resulting funds will be used effectively. There is a vast chasm between true planning and "grantsmanship." Schools are definitely shortchanged to the extent that the evaluation and planning involved in the grantwriting exercise are "simulated" rather than "real." But if the evaluation/planning effort is real, a well-developed written plan format can be of immense value in insuring that the fruits of planning are not lost but remain available for use during the implementation phase.

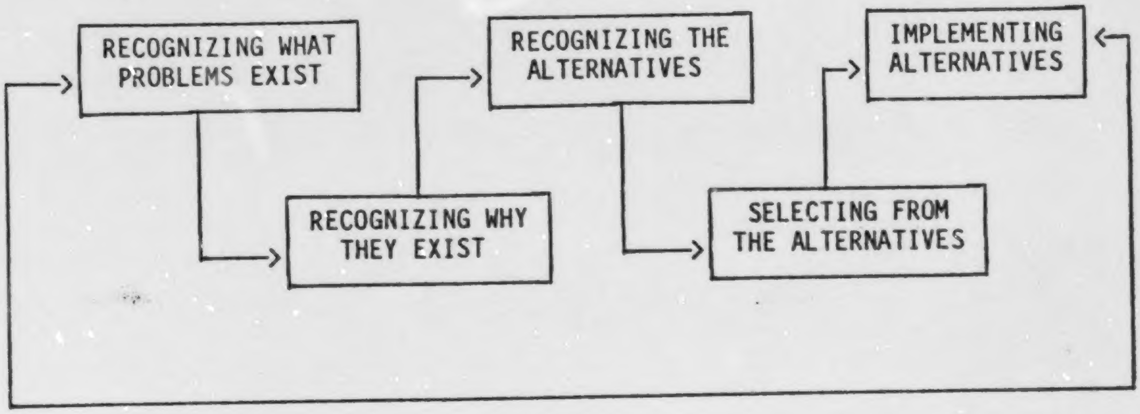
You will notice we talk about the concepts of evaluation and planning as inseparable companions. It is our belief that the value accruing to the school district relates very directly to the closeness of the linkage between the two. In fact, we believe that the major reason for evaluation is, in fact, planning. We believe that evaluation should exist to enlighten decision makers.

The close relationship between evaluation and planning depends upon an equally close relationship between evaluator and planner. Within our district, this problem is easily solved: they are one and the same person. The Evaluation Specialist is also the Planning Specialist, though not by name. To some evaluators this may be a scandalous situation. It is, therefore, important that you understand our position. In our district, most evaluation and planning activities take place at the school site. The Evaluation Specialist acts primarily as a resource person and facilitator. Most of the actual evaluation and planning "conclusions" are reached by those participating at the local level. The Evaluation Specialist is the developer and custodian of the evaluation and planning processes. Those at the local level are the actual participants. The Evaluation Specialist is also the manufacturer and provider of data. Those involved at the school level are the actual consumers. For these reasons, much of the traditional concern for evaluator independence is unnecessary. While the Evaluation Specialist could certainly be expected to be biased with regard to the value of the processes involved, he does not necessarily have to be biased with regard to the program which results at any particular school site. What this dual role does do is prevent the impotence which often results when the evaluator is absent during the critical latter half of the evaluation/planning process.

Our Planning/Evaluation Cycle

We have, thus far, discussed our philosophy as to the "why's" of evaluation and planning. We have thoughts regarding the "how's" of evaluation and planning.

First, any evaluation/planning cycle should parallel the essential steps of human problem solving which are. . .


RECOGNIZING WHAT PROBLEMS EXIST
RECOGNIZING WHY THEY EXIST
RECOGNIZING THE ALTERNATIVES
SELECTING FROM THE ALTERNATIVES
IMPLEMENTING ALTERNATIVES

Our evaluation/planning process acts as our "tour guide" leading us through each stage of problem solving.

Our evaluation/planning process is comprehensive, in the same way a physical examination is comprehensive. The medical doctor cannot affort to overlook a single area of medical significance to his patient. Nor can we, as educators, afford to overlook any of the educational essentials if we wish to maximize our students' educational wellbeing.

Our evaluation/planning process is "broadbased." It involves all elements of the school community. This includes students (at the proper level of maturity), parents, staff members as well as administrators. The value of the varying perceptions and points of view which exist within the various groups should not be underestimated. Also, being participants in the processes of evaluating and planning can develop a deep sense of commitment toward the resulting program as well as a kind of "bonding" between school and community. This participatory aspect of evaluation and planning places the site administrators in a unique and key role.

Our evaluation/planning process is "data-based." We mentioned that we believed it was evaluation's primary responsibility to enlighten planners or decision makers. One means of enlightenment is through available information regarding student performance, attitudes, etc. We have contacted and "heard from" each student through the medium of a written instrument. We would seldom, if ever, have the time to do this in any more literal way. So, surveys and the like serve us by gathering those student, parent and staff perceptions which we acclaimed earlier. The more such information we have, the better our decisions are likely to be. Our educational program and our students are the ultimate beneficiaries.

The Evolution of Our Model

One has to reflect back over ten or fifteen years to recognize the dramatic changes which have occurred in the way we operate our schools in the area of evaluation and planning.

Initial Resistance

At the beginning, my feeling was one of loneliness, like being out on a limb all by yourself. The attitude in our schools regarding "outsiders" -- not only local visitors, but also staff members and administrators from other schools -- was often pleasant but guarded. The feeling toward central administrators was often one of privacy, protection or suspiciousness. When we began getting down to business, the atmosphere frequently became even less pleasant. And the differences were not limited to attitudes.

The aversion to what might be described as program management activities seemed widespread: "Principals are teachers' teachers, not managers!" Since formalized planning was much less frequent than it is today, programs were based on enduring, but perhaps obsolete, assumptions rather than on any assessment of participant needs. Programs were incompletely planned, with significant areas regarding objectives and curriculum omitted. There was only a limited sense of accountability for program funds regarding such programs as ESEA Title I. Staff commitment to such programs was limited by a general feeling that these programs had been "laid on." Lastly, but very significantly, there was limited parental support for such programs frequently accompanied by criticism and suspicion.

I mentioned that staffs felt a limited sense of commitment and also that programs were "laid on." Such was exactly the case, and so specified by state regulations. At that time, districts throughout the state centrally developed a single Title I plan and then shipped it out to the project schools for implementation. The program contained objectives which served as the basis for a single central program evaluation. These objectives were not considered by the schools to be their objectives.

It was this condition that provided the impetus for change to the site-level approach we see in use today.

Ten years ago, when our Evaluation Specialist position was established, the district had six elementary schools in its Title I program. School personnel continually asked that they be allowed to create their own program plans suited to the uniquenesses of their individual school sites. Serious doubt was often expressed by federal or state administrators with regard to the ability of local school sites to accomplish such a task. However, in 1971, the legislature passed legislation creating the Early Childhood Education (ECE) program and, with it, a framework for evaluation and planning at the school level. The California State Department of Education developed the evaluation/planning model as well as the planning format. The creation of the ECE program prompted us to develop a participatory evaluation/planning model based on a broadbased needs assessment process. Our efforts were shaped by the state's process and by its forms.

Before discussing what we did, I think it would be worthwhile to discuss the atmosphere within the "central office" toward our efforts. Early on, central administrative awareness, understanding and appreciation of our efforts were limited as was their support. There were several highly placed strategic exceptions, namely, the Director of Research and Evaluation and his superior, the Assistant Superintendent for Special Services. They have always philosophically supported the importance of educational evaluation.

We also had a critical ally in the ECE/Title I Director. He had based his doctoral dissertation on community needs assessment. Many times, he and I shared the feelings of isolation which were so much a part of those early days. At one point, he asked me if I realized that it was just he and I trying to convince a skeptical district of the value of categorical programs, planning, evaluation and the like. I will never forget the "friendly" weekend mountain retreat for project principals which evolved into a kind of marathon Director/Evaluator "roast" lasting two days.

We knew that for our efforts to be successful we would have to win the central and field administration over to our point of view. We attempted to accomplish this by deed. We know that evaluation and planning activities do create a burden for teachers and principals. Our evaluation/planning activities had to solve more problems than they created. The values of our processes had to be made visable to students and management. We successfully used our model several times to facilitate the resolution of school-community conflicts in schools outside the small group of project schools. We can now report that our relationship with both central and field administrators is an exceedingly positive one. It is satisfying to see, and especially to feel, the contrast.

Needs Assessment Survey

Let's return now to our needs assessment process. Our earliest versions of the staff, parent and student survey instruments came to be widely known as the Educational Program Assessment (EPA) process. Since its creation, the ongoing work on the EPA instruments has involved large numbers of parents, staff, administrators and secondary students. This has given us a unique opportunity to test out a theory regarding the beneficial effects of involvement on school and community feelings of ownership and support for such processes.

We believe that broad school and community involvement has contributed to responsive instruments. These, in turn, have led to high response rates, averaging about 70 percent and rising to an unbelievable high of 100 percent response from one school and its community.

There were, I believe, several additional key elements in our success. First, our surveys were very unusual in that they, unlike many others conducted in the educational setting, were done "by" the school and "for" the school. It was one of those rare instances where information actually completed the loop back to its point of origin. School people were used to seeing a survey when filling it out and never again. This time the resulting data was returned to them for their use, a very significant difference.

While school-site involvement in instrument development and completion of the information loop were necessary, they were not sufficient to ensure the positive effect we wanted from our efforts. We carefully considered the content, the format and the length of the instruments themselves as well as to the manner in which the resulting data were presented for use at the school level.

The questions in our survey instruments addressed not only those areas of particular interest to the sponsoring special program(s), but also those areas of general concern to the entire educational program. In this way, much of what a school wanted to know from its community could be obtained from a single survey. Survey questions asked about parents' perceptions of the effectiveness of the current program in various areas of the curriculum as well as the need for improvement in their child's performance. Questions were limited to those which respondents could reasonably and intelligently answer. For example, parents were asked about the performance attitudes, etc. of their own children and not those of other students in the school or neighborhood.

As I mentioned, we were concerned about the way in which the data was delivered back to the school. Here, we ran into a dilemma. Popular at that time was a belief that broad participation in any activity, including survey tallying, was beneficial. This is precisely how we formatted our surveys during the first year. This was, in part, to test our belief in participation but mainly it was practical. There was no other way of doing it at the time. All went well, but it became apparent to me that the cost, in terms of time, accuracy and effort, exceeded the actual benefits. I came to believe that the greatest benefit accrues from easy access to the resulting information, not from the process used to analyze it.

The second set of arrangements made for survey processing was even more interesting. We used the computer facilities of a university in a neighboring community, having keypunched the surveys within the district. The data decks, along with the control cards for the statistical software being used had to be physically transported to the university computer center. This I usually did on my way home from work. The next morning, I retrieved the decks and printouts on my way to work where a team of clerical workers transcribed their contents onto a specially developed graphic format. In spite of its compexity, this arrangement was a definite improvement over hand tallying.

But finally, we moved into a third phase for survey processing. My commuting back and forth between the district and the university for one year provided the district with the incentive to develop our own computer program for use within the district. We then could tabulate the results of each survey in both graphic and tabular form.

Standardized Tests

While survey responses do provide a large amount of information for evaluation and planning, they do not provide it all. The primary source of our evidence about student achievement is the standardized test, in our case the Iowa Test of Basic Skills (ITBS). We needed to maximize the value and validity of the test information while minimizing its cost in time and effort. Title I regulations had mandated fall and spring testing which seemed an enormous amount of testing with its predictable negative effect upon the students and staff. Calculating academic gains between fall and spring, as was the practice, brought with it all sorts of complications and excluded the significant effects of the summer period. Ironically, due to this spring testing requirement, the most useful information for local evaluation and planning typically became available to schools at or after the end of the school year. For these reasons, the districts in the Sacramento area requested, and were granted, a waiver so that they could implement an annual (spring-to-spring) testing schedule. The gains from such testing cumulate over a number of years and yield a more accurate reflection of student learning. Perhaps the greatest benefit was that the information for evaluation and planning was now available at the outset of the next school year.

While we had overcome a major problem in the area of standardized testing, another related major problem went unattended until very recently. This had to do with the form in which standardized test results were provited to schools. This was not a problem unique to the San Juan Unified School District but can be found coast to coast. District offices, as do the scoring services offered by test publishers, ordinarily return results as computer printouts, usually perceived by teachers as unattractive and technical. However, this is only half of the problem. The other half has to do with the staggering number of such printouts which must be consulted in order to draw the kinds of conclusions that are critical to evaluation and planning.

For example, a test with four subtest areas will require 24 printouts to summarize a single year's results (grades 1-6). The typical Title I school will also want separate results for participating students, requiring another 24 printouts. Much of the value from standardized test results comes from the examination of results over a number of years. Looking at a four-year period for these two groups requires 192 pages of printouts. For a separate report of gains for each of these groups an additional 120 pages is needed. That is a grand total of 312 pages, more than enough to intimidate the most ambitious administrator.

This problem led to the development of our Academic Performance Profile (APP). Each Academic Performance Profile contained graphic and statistic results for a particular subtest area. On this single profile are the results for "all the grades" for "all the years." A set of eight profiles contains the information formerly contained in 312 pages.

Another of our conquests is a major advance in the evaluation/planning process, and deals with the actual act of budgeting. Formerly, it was left until after planning was complete. But, our recent modification to the planning format has created a proper place for budgeting with in the evaluation/planning process. Each activity in the program description is now linked to the funds and the funding source which makes it possible. An easy-to-read "audit trail" has thus been created.

An evolutionary process has, thus, brought us to a point where we have a needs assessment process, and an efficient and effective means by which to present standardized test results.

A Description of Our Model

Any discussion of our current evaluation/planning model should focus on four major areas: our human evaluation/planning activities at the school site, our program plan forms, the information gathering and organizing processes and our district level evaluation/planning activities.

Evaluation and Planning at Schools

As we mentioned earlier, the evaluation/planning model is systematic problem solving, i.e., identifying existing problems, isolating the mental reasons for those problems, discovering and evaluating alternative solutions, adopting the most promising solutions, and going about our business. Our evaluation/planning process takes people through six steps which gets them to problem solve at the school-community level.

Evaluation and planning activities at the school take place under the patronage of the School Site Council, a body whose composition and influence has grown out of the California School Improvement Program (SIP) legislation. This committee, with the site administrator, represents each of the major constituencies within school-community (i.e., parents, staff, administrators, and secondary students).

The school's curriculum is typically (although not always) separated into academic components such as reading, language, mathematics, etc. The School Site Council usually creates "component committees" of parents, staff members and students (in secondary schools) to do evaluation and planning for each academic area. The final school program plan will have sections addressing each component written by the separate committees.

At the outset, each component commitee tries to define their area of the curriculum comprehensively, to ensure that all sub-areas are addressed in the evaluation/planning process.

The committees of educational planners, like physicians, begin by determining the status of their students' educational "health" in each significant area.

When a physician's examination and tests are complete, his or her job has only just begun. The same is true for the component committees. The members of the component committees examine their newly gathered information to identify the strengths and weaknesses of their students. With this step complete, the component committees turn their attention to the existing programs. In those areas where student performance has been found to be unsatisfactory, an effort is made to identify probable program causes, that is, program gaps or weaknesses which, if eliminated, would result in improved student performance.

Once a doctor has completed his review of the existing treatment plan, he is then ready to progress to the presciption phase, in which a new, or perhaps a modified, treatment plan is developed. Not unlike the physician, educators have the responsibility of remaining aware of the variety of strategies and materials which are potentially available to meet the needs of their "clients". School people join with others who have special skills in the curricular areas and are available for consultation. After such consultation, members can modify, extend and improve their school's existing program. When they have finished deliberating, they use the Program Plan form to document their new program description.

Were schools not constrained by budget and time limitations, it would be simple to adopt all of the modifications and additions proposed by the various component committees. However, dollars and time are definitely limited. Therefore, there must be a forum to make these inevitable time and financial budgeting decisions. We return to the School Site Council who does a component-by-component review of the proposed program after an anticipated projected cost has been attached to each proposed expenditure. Through negotiation and compromise, the activities of lowest priority, along with their related costs, will be reduced or eliminated to the point where the amount which the school proposes to spend will exactly equal the amount which it expects to receive.

There is something of great importance which occurs at this critical juncture in the evaluation/planning process. Up to this point all activities have occurred in the individual component committees. In order for ownership and commitment to be strengthened there must be this meeting of the members of the total planning group. The ultimate result, sometimes after many emotional hours, is a program which is widely owned and supported.

When the doctor has completed his or her diagnostic activities and prescribed a course of treatment, he is often in a position to offer an opinion regarding the probable effects of the treatment plan on the future health of the patient. This is called a "prognosis." Our schools accomplish this step when they identify their program objectives and determine a means of measurement for each. It is our belief that the greatest value of objectives does not come from identifying them but rather from the data gathering activities which they stimulate. It is also our belief that objectives become much less fanciful when they are not prematurely formulated.

Program Plan Forms

We have already mentioned that our program plan format was designed so that there is a very close relationship between its layout and the various steps of the evaluation/planning process. The columns are sequentially organized from left to right to parallel the steps of the evaluation/planning process. The format assists the experienced planner/evaluator in documenting his/her findings. It functions as a guide for the less experienced by presenting the steps in their proper order.

This format has undergone a number of major revisions. The State's first attempt at a form for use was a complicated two-page affair with needless redundancy. The State's willingness to allow districts to experiment with the form itself has led to major improvements.

Information Gathering and Organization

We have described the survey instrument on which we gather the views of each of the three major segments of the school community -- parents, staff, and secondary students. The survey is composed of a minimum set of standard questions common to all schools plus a second set of optional school-developed questions unique to each school. Our assistance is provided to requesting schools in developing questions. The district summarizes these school-developed questions along with the standard questions used at every site.

Each staff member and secondary student completes an EPA survey form, while the parent survey is completed one-per-household. The responding parent is directed to randomly select one child as the survey's focus.

One set of "standard" survey questions request that the parent or secondary student assess the effectiveness of the school in teaching the student what he/she needs to know in each of fourteen curricular areas (eighteen at the secondary level). This set of questions tries to establish the status quo in terms of the extent to which the student's academic needs are being met. There is another side to the coin: how much or how little need is there for change in the status quo? That is, how much need is there for improvement in the student's performance in each of those same areas? A second "standard" set of questions brings student and parental values into play. The fact that a particular student's needs are not being met effectively in a particular area of the curriculum does not necessarily mean that program improvement is desired either by the parent or the student. The two sets of questions must, in our view, be used together to produce a complete picture.

Surveying takes place in the spring of each year. Our office schedules evaluation/planning release time. We are aware of when each site's surveying must begin in order that the tabulated results are available during their scheduled release time. The standard survey forms are printed and shipped from the central office. Any desired "optional" school questions are added at the school prior to their distribution. The surveys are returned to the central office for processing.

Earlier we described the exhaustive efforts that have gone into the presentation of survey and other results. Currently, the data from EPA survey instruments are keypunched directly onto tape which is then formulated into tabular and histogram formats. The survey uses four-point "forced choice" questions in order to enable the creation of a simple visual index of, for example, "effectiveness." Almost instantly, the strongest and weakest areas can be identified by consulting the histograms provided.

While the student and parent survey instruments focus almost entirely upon student performance, the staff survey asks for comment on the strengths and weaknesses of key educational activities. This has been most extensively developed in the secondary school version of the EPA staff survey where clusters of questions address topics such as the effectiveness of the instructional processes, the effectiveness of the curriculum content and organization, the effectiveness of the class and school environments in which learning takes place, the effectiveness of staff development, the effectiveness of guidance and health services. Responses to each of these questions are summarized and presented on a school-wide departmental basis. We are developing similar questions for the elementary staff survey instrument. The same computer program is used to summarize these items as well as the other sections of the staff survey. It should be emphasized that this computer program was written so as to be useful in summarizing a wide variety of local evaluation instruments.

While the surveys which make up the EPA are a valuable source of evaluation and planning information, they by no means provide all of the necessary data. Academic performance in our district is assessed by the Iowa Test of Basic Skills (ITBS). One major product of the recent past was the development of an economical and concise method of storing and presenting the results of the ITBS. The method developed was the Academic Performance Profile (APP). The APP conveys a large amount of information relating to the academic performance of students assessed by the ITBS. Admittedly, it has a complicated appearance at first glance. However, once understood, it is of great value in enabling the reader to master a lot of information in a minimum amount of time.

San Juan Unified School District annually assesses the academic performance of its elementary and intermediate students during the month of January. Our profile format relies on grade equivalent scores, although the format can easily be modified for use with percentile scores.

Each profile corresponds to a single ITBS subtest. All ITBS grade equivalent information pertaining to the particular subtest is contained in that profile regardless of the grade or year. The basic information contained in the Profile is displayed in two ways: 1) in statistical fashion in tables situated down the right side of the Profile; and 2) in visual or graphic fashion in the graphs down the center of the Profile.

A key table located at the upper right corner of the Profile contains grade placement scores for each grade and year. These scores can be examined by grade level (horizontally) or by student group (diagonally). These two types of analyses are shown via the top two graphs in the center of the form. Across the bottom of the form, a single graph and table present year-to-year group gains for the various grades in an unusual manner. The gains are stacked, one atop the other, to produce an aggregate indicator of the "total productivity" of the school.

Within our district, the availability of the ITBS data in this fashion has greatly expanded its use. A similar profile has also been developed to display the standardized test data for Title I participants only. This profile, in contrast, charts the beginning and ending grade placement scores as well as the related grade placement gains for all students identified as Title I participants.

District Level Evaluation and Planning

Each of our evaluative instruments as well as their summary displays is designed to satisfy district and program evaluation and planning needs. Responsive to the informational needs of the board of education, district and program administrators, such summaries also serve as valuable frames of reference for those at the local sites. In many instances, these district-level summaries are placed on transparencies directly laid over the school's summary for easy comparison.

In all of these activities, our motive has been to create the greatest benefit with the least burden. School staff members must be involved in program evaluation and planning. They also have a fundamental responsibility to be in classrooms, educating students. Therefore, their time out of the classroom must be well-utilized. We must see that the benefits to students outweigh the cost of administrator and teacher time.

A Look To The Future

While we are proud of our accomplishments to date, there are always areas in need of attention and improvement. We do not allow ourselves the luxury of feeling totally adequate. We know that we have a long way to go. There are many areas where it still "just doesn't feel right." In some instances we know why; in others, we do not.

In this age of tenuous financing throughout the educational system, perhaps a greater proportion of the potential value from evaluation and planning activities could be had if something could be done to consolidate special program funding. But this is not the major problem. While the size of the check is of importance, a great deal of difficulty is caused by the late date at which districts and, therefore, schools find out what their actual funding is going to be for the coming year. Making the most efficient use of our limited financial resources is one of the most basic goals of evaluation and planning. Its attainment is not made any easier when we find out what our budget is only after our planning efforts have been concluded. This complication in our life has existed for a long time, probably as long as the special program funds have existed. It should be mentioned, because it is one of the factors limiting our success in evaluation and planning.

While we are speaking of budgets and budgeting, there is another situation worth mentioning. There is a definite need for inservice training in the area of accounting and bookkeeping for staff members and administrators alike. Special program funds and accountability for those funds, not to mention the budgeting act itself, have been a source of frustration to those inexperienced in such matters. We have only scratched the surface of automated expenditure accounting in the management of special program funds. The typical district accounting system is not equipped to handle the detailed budgeting required by special programs. Participatory decision making requires up-to-date figures regarding expenditures, unencumbered reserves and the like. The application of modern microcomputer and word-processing technology holds great promise.

Meeting all of the widely varying demands for information at the school, program, and district level is a great challenge. Automation is, I believe, the ultimate answer. Few people fully appreciate the magnitude of fully automating a district's data needs. Computer programming is a slow and arduous process in the beginning, especially when there is intense competition for computer services. Shrinking funds complicate matters further. We have to look into the future for such advances as an accessible district-level data base. When this is a reality, evaluation and planning will realize more of their potential. Theoretically, there are no limits to the factors that could be encoded in a data system. Entering a reasonable number of "independent variables" could enable us to select one or a combination of these variables and examine their longitudinal effects.

Our goal must continue to be to maximize the benefits while minimizing the burdens of any activity outside the classroom, including evaluation and planning. While our district has been very successful in recent years in reducing the burden, there is stil room for improvement. We must continue to emphasize bona fide planning as opposed to "simulated" planning or plan writing. We intend to focus our attention on increasing the depth at which we examine our instructional program, concentrating on effective instructive strategies. We have come a long way. We still have a long way to go.

CINCINNATI'S EXPERIENCE: SUPPORTING COMMUNITY EFFORTS TO IMPROVE INSTRUCTION: A FOCUS FOR EVALUATION

by Joseph Felix

Introduction

Improving instruction despite declining resources is a key challenge facing public education today, especially in urban districts. School finance has become the crucial issue, debated by educators, politicians, parents, taxpayers and almost everyone else. The financial difficulties that have forced many schools to revise their curricular offerings, limit after-school activities, and even close their doors at times are not expected to disappear in the near future. In fact, with a continuing enrollment decline and an aging population with fewer direct ties to schools, the picture may even become gloomier.

For many school districts across the country, budgetary austerity has become a way of life. California's Proposition 13 is probably the most publicized instance of taxpayer revolt that has cut into schools' budgets. But the problem is nationwide, especially among large-city school systems.

Ohio's requirement that each increase of school taxes be approved by the voters has made the problem particularly acute for our large cities. Columbus, for example, has had to make dramatic cutbacks in its research and evaluation area, as well as in other departments that had been built to notable strength. Cleveland has been on the brink of bankruptcy.

In Cincinnati, too, the financial problems loom great. A $12 million deficit has been projected for the 1983 fiscal year. Confronted with such fiscal shortage, it is very difficult to maintain formal efforts to improve teaching and learning.

Some years ago I co-authored an article for Theory into Practice on Research and Development (R&D) in Cincinnati. I gave the article the alliterative title, "Developing Developers: The Race to Improve Education." It began:

> Read that title again. Do you hear hoofbeats?
>
> The race is on . . . For big-city liveries . . .the odds against bringing home a winner are astronomical. Undernourished by voters' decree, our horses move forward at a pace that would humiliate the tortoise. A seemingly endless parade of jockeys mount and dismount with unprecedented frequency.
>
> Cincinnati's chances of winning the race are probably not much better than those of any other major city. But we've scouted up an entry fee from our benevolent Uncle Sam, hired a few stable boys and girls to unstabalize things, and mounted a drive for change . . ."(1)

When R&D was given its start through federal funding, there was much uncertainty about how long it would last. "Soft money" was seen as an unsound foundation for so important an endeavor. We made efforts to get R&D services moved to the "more solid ground" of local tax money. Now the hard ground is even softer than the soft ground. No wonder our horses stumble occasionally!

Gratefully, I have seen a strong R&D effort survive in Cincinnati despite severe budgetary limitations. Credit for this survival must go to our Superintendent, James N. Jacobs, who has national renown as a leader in public school research and evaluation. Jacobs had won the confidence of the board and much of the staff and community for the cost-effectiveness of

---

1. Jacobs, James N., & Felix, Joseph L. "Developing Developers: The Race to Improve Education." Theory Into Practice, April, 1972, 11, 225-231.

evaluation and program development. We continue to believe that a strong evaluation unit can lead a school system to making the best possible use of its limited resources.

Recently, we have developed in Cincinnati a major thrust that holds considerable promise for further strengthening our instructional program. The board and administration have been making a concerted effort to increase community involvement in both the decision-making and support processes. I want to describe this move to bring the community closer to the schools and the measures we have taken to make community involvement the focal point of our evaluation effort.

Despite our budgetary problems, the Cincinnati school system is at an historic juncture. Even as local, state, and federal funding decreases, the interest of the community in contributing its time and energy for the support of schools continues to rise. The business community has rallied over the past five years to furnish us with substantial free consultation and support services. A number of local businesses have established partnerships with individual schools and are contributing both time and money to the improvement of these schools. Volunteer service throughout the district is increasing in the face of a contrary national trend of decreased voluntary contribution of time.

In addition, cooperative efforts have begun with local universities, the City of Cincinnati, and other political and educational groups. These endeavors hold considerable promise for filling the gaps caused by limited financial resources.

What has brought about the school district's success in garnering community support? The contributing factors are numerous, but an essential ingredient has been the school district's willingness to allow participation in the decision process. Our attitude creates, in turn, a very positive attitude among the participants.

Of course, organizations such as the PTA have long contributed time and resources to helping us attain school goals. Their willingness and interest in investing in the goals set by the board and administration might have dwindled in Cincinnati as elsewhere. But by inviting parents and other community members to participate in planning, our school system is gaining commitment to the educational program. We know that the community's willingness to contribute volunteer resources and to approve tax support is increasing. A new organization called Cincinnatians Active To Support Education rallied behind a tax referendum three years ago and gained its approval. Community councils are beginning to take an active interest in education, and this often occurs in neighborhoods where other forms of parent and community participation are minimal.

In 1980, the district started a new program of community involvement. This program is a comprehensive, decentralized educational planning process carried out by teams representative of each school's total community. Because of the district's financial crisis, the initial focus was on budgetary planning. The program was thus designated Local School Budgeting (LSB). More recently, the Board of Education passed an action motion calling for the establishment of local school advisory committees in all schools. The budgeting teams are seen as the prototypes of these advisory committees.

As with any new programmatic effort, initiating LSB has not been problem-free. Community participation in decision making is threatening to some administrators. Reaching consensus on needs and goals is arduous when different segments of the community have varying priorities. Many questions of prerogative and process have had to be resolved as the LSB program has moved forward. Still, we see the program as a significantly successful means of improving the education we are providing for the students of our district.

We have chosen to make the local school planning effort the focus of our evaluation services. We define evaluation in the Cincinnati schools in the broad sense of providing information for decision making. For the past ten years, the structure and functions of the Evaluation Branch have shifted towards serving the local school unit in preference to serving central administrators. In the hope that managers in other districts might benefit from our experience, I would like to describe how Cincinnati's evaluation unit has attempted to respond to the needs of school-community participants in the educational planning process.

To help the reader understand the setting in which we function, let's look first at the demography of our school district and the history of its evaluation unit. Then we'll discuss Local School Budgeting and the branch functions that support it. Finally, we'll talk about the specific role of the Local School Evaluator.

The Cincinnati School District and Our Evaluation Past

The Cincinnati school district covers an area of 90 square miles with a population of 410,000. The school district's boundaries overlap a number of politically separate areas. In addition, there are within the circumference of the district several communities that have their own government and school district.

The enrollment of the Cincinnati Public Schools presently totals approximately 51,000. Of these, 57 percent are Black, 42 percent are Caucasian, and one percent are of other races. The socio-economic level of the district is reflected in the fact that 56 percent of the enrolled students qualify for free or reduced-price lunches.

There are 86 schools in the district: 57 elementary, 13 middle or junior high schools, 9 senior highs and 7 special schools for handicapped students. To achieve racial balance, the school system operates a voluntary integration program featuring a wide variety of alternative curricula and an open enrollment policy.

In 1982-83 the district had a general fund operating budget of $138 million. This was supplemented with about $8 million of federal grant support and $3 million of special state funding.

The foundations of research and evaluation in Cincinnati, as in most other school systems, are in the testing and survey movements of the 1920's and 1930's. Prior to this time, inquiry and innovation rested almost entirely in the hands of the creative teacher and administrator. In 1929, the Division of Research and Statistics was organized in Cincinnati Public Schools. Nine years later, a separate division in the Department of Instruction was set up to administer the group testing program. These two divisions provided the first semblance of systematic structure and scientific orientation for R&E in Cincinnati.

With the passage of the Elementary and Secondary Education Act, there came a new emphasis on program evaluation. The mandate that the federally funded projects developed under Title I be systematically evaluated led to the creation of the Division of Program Development in 1965.

Since that time, research and evaluation activities have been conducted by personnel in a variety of organizational structures. For several years, the district had a Department of Research and Development headed by our current Superintendent, then an assistant superintendent. Jacobs was named Interim Superintendent of the Cincinnati Public Schools in July, 1976. Regular appointment as Superintendent of Schools followed in February, 1977.

As the district's financial condition worsened, some cutbacks in R&D -- as everywhere else -- were essential. With much pain, Jacobs disbanded the department he had organized. Branches that had been part of the department were assigned to other departments. Functions were kept intact.

Throughout the recent organizational changes, the Evaluation Branch has remained responsible for most R&E activities. Changes in the organizational structure have made it easier to coordinate some functions such as test administration and interpretation. But these same changes have also added difficulty and challenge to the task of integrating evalution with program planning.

The Evaluation Branch includes four sections: Program Evaluation, Testing, School Information, and Communications. These four sections cooperate to serve district and school needs. Increasingly, however, a higher priority has been given to meeting the needs of the local school units in the district. In discussing this process, I'll try to show how the four sections of the branch relate to one another and to the local school planning effort.

Early evaluations of Title I projects were organized to meet state and federal requirements. To receive continued funding, the school district had to complete state evaluation forms and submit written reports of results. Local interest in the findings thus submitted was limited, and evaluation outcomes did little to impact program development or administration.

After several years, questions began to be asked nationally about the effectiveness of the Title I investment. This brought the local district increased pressure for improvement in these programs. Gradually, it became common to design evaluations so they served school district decision-making, as well as state and national reporting needs. Still, the perennial conflict between scientific results and political expedience continued to limit the benefit derived from the evaluation effort.

Somewhere around the beginning of the 1970's, our emphasis began to change. Individual schools in our district were seeking autonomy in planning and implementing their instructional programs. The school district obtained a grant under Title III ESEA to create a "School-Community Evaluation and Development System." This system set up within one senior high school area a working partnership between school and community to develop and evaluate instructional programs. From this project there emerged a model for local school program development and evaluation which, although never officially adopted by the Board, has strongly influenced school district evaluation services. (2)

---

2. Jacobs, James N. "A Model for Program Development and Evaluation." Theory Into Practice, February, 1974, 13, 15-21.

Local School Budgeting (LSB)

LSB has been in Cincinnati for just over two years. Already we believe it has altered the traditional modes of educational planning. Although evidence is still scanty, our current data suggest improved student achievement in those schools where the program is working well. We also believe that these same schools are typically directing their educational resources more pointedly at their goals. If the evaluation effort can meet its challenge of adequately supporting this movement, the program holds great promise for getting more mileage from the educational dollar.

To initiate the program, a Local School Budgeting Steering Committee was chosen by the community. This committee selected seven schools to participate in the first year. Three other schools were added to the program through a foundation grant. Additional schools are being included each year so that every school will be participating by 1985.

Creative approaches to resource allocation is the basic challenge confronting LSB teams. The project has been designated as a "budgeting" project to focus on this need. More accurately, it is a decentralized process of educational planning that attempts to involve the total school community of each local school.

There are numerous examples of creative use of resources devised by the budgeting teams. Several schools have elected to organize volunteers to supply some of the school's maintenance needs, thereby decreasing the janitorial budget and providing more funds for instruction. Schools have formed partnerships with local businesses and industry. They have organized pools of qualified volunteers to serve as substitute teachers so that money formerly spent for substitutes' salaries could be dedicated to other purposes.

One of the most common steps taken by school budget teams has been the application of energy-saving measures. Savings realized through such economies are returned to the school for instructional use. Other schools have capitalized on the school district's current practice of providing paint for volunteers to redecorate classrooms and corridors when budgetary limitations preclude having salaried personnel perform this service.

More related to evaluation activities are the teams' efforts to strengthen the instructional process. Each team has an opportunity to locate personnel, material or training resources for this purpose. When a local-site budgeting team attempts to stretch its resources to meet the needs of students in the school, it is dealing with much more than dollars. The team needs to have some way to choose between alternative means of meeting needs; evaluation of possible programs is important. Techniques for program planning in times of austerity are also important to make the best use of limited resources. The team also needs to know of ways of soliciting funds from external sources, building volunteer support, and capitalizing on community and parent resources. Training and consultation on these and other topics are available to all team members.

To better serve the needs of students, the program locates decisions at the individual school. This decentralized process takes advantage of the day-to-day contact between teachers and students and uses that interchange to add valuable knowledge to decision making. Non-certificated staff members and members of the community also apply their experience and knowledge when assessing the strengths and needs of their local schools.

The rationale for LSB recognizes that parents are also prime teachers. They have specific concerns pertaining to their children's education. They are also uniquely aware of their children's needs. Since the public at large supplies the resources for education, they also have a legitimate interest in what is being done with those resources.

Team members are selected by their respective groups: parents elect parents, community council members and staff choose their representatives. The principal, as the accountable educational leader of the school, serves as a member of the team, but also acts as the decision maker on many local issues.

Team members bring ideas to the committee and reflect the position of the group they are representing. This approach brings non-team members closer to the decision-making process. It enables a parent to contact a person on the committee and have his/her opinion expressed.

The underlying conceptual model for evaluating services throughout the system is Stufflebeam's Context, Input, Process, Product (CIPP) Model. (3) This model sets forth four stages for evaluation as a means of "delineating, obtaining and providing information for decision making." The first stage is context evaluation, which involves needs assessment and goal development. Input evaluation, the second stage, includes the review of various

---

3. Stufflebeam, D. L., et al. Educational Evaluation and Decision Making. Ithaca, Illinois: Peacock Publishers, Inc., 1971.

alternatives that might accomplish the goals that have been selected. In the third stage, process evaluation, the implementation of the selected program is monitored to insure that plans are carried out according to schedule. Finally, product evaluation looks at the outcomes of the educational program to determine whether objectives and goals have in fact been attained.

The LSB planning model has been designed to fit the CIPP approach to evaluation. It also includes four stages: determining needs and goals, resource allocation, implementation, and review.

Key support for the local teams is provided by the Evaluation Branch; the Communications section coordinates this support. This coordination includes training and consultation. In addition, a parent-community coordinator helps principals and teams build local interest. Communications specialists garner media support and publish information to facilitate the work of the teams.

Other sections of the branch provide support services more traditionally considered evaluation functions. These services are reflected in the discussion of the four stages of the planning model.

Needs and Goals. Assessing needs and setting goals are continuing processes. Formally, however, local needs assessment begins in January of each year. A goal development model which we evolved from a former federal project is recommended to each local team. The School Information Section provides crucial service in this stage of planning. A data bank called the School Information System includes several hundred variables on each school unit. A number of these -- e.g., dropout rates by cause, socioeconomic variables, etc. -- are very helpful in determining high-priority local needs. Also in the data bank are the results of annual surveys. Each year surveys of administrators, teachers, parents, and students are conducted throughout the school system. Results of these surveys are useful barometers of attitudes and opinions.

Resource Allocation. Once goals have been set, the team next addresses the allocation of resources. Monies allocated to the local level are considered first; the team determines the best use of these dollars in attaining the goals that have been set. Resource allocation goes beyond the available dollars, however, as teams then address other ways to met their goals. Plans for obtaining external funds, using community resources, building volunteer resources, and teaching parents to work with their children toward school goals are all part of this stage.

In addition to training, grantsmanship assistance, and community relations help, the Evaluation Branch provides support in prioritizing and evaluating possible program strategies. A federally funded Cost Effectiveness project has generated a model by which teams may study alternative program possibilities. A library research capability within the School Information Section furnishes guidance from the professional literature.

The resource allocation stage of the model is seen as the focal point for evaluation support. From the team's evaluation in this stage flow recommendations and decisions that determine the local program.

Together with the implementation (next stage), these functions constitute the power cycle that lies at the heart of LSB success: Evaluation, Recommendation, Decision, Implementation. The LSB philosophy recognizes that power resides in shared participation in these functions. It minimizes the importance of questions like "Who makes the final decision?"

The final local school budget for the coming year is developed by the budget team, then presented to the public at open hearings at each school. After public feedback has been taken into consideration, the final budget is given to the central administration for approval. If the budget is unacceptable, the administration discusses the problems with the budget team and together they resolve the issues. Finally, the Cincinnati Board of Education adopts the annual district budget reflecting the individual school needs.

Implementation and Review. The two remaining stages of the planning model are implementation and review. The evaluation support for these stages follows the traditional modes of process and product evaluation. The Program Evaluation section, using data from the Testing and School Information sections, furnishes this support.

Implementation and review data are made available by several additional evaluation-related activities. The standardized testing program provides survey achievement information in reading and mathematics, grades one through eight. The Program Evaluation section has devised several creative ways to tie this information more directly to instruction.

Other activities supporting implementation and review are a program of criterion-referenced instructional management called the Cincinnati Instructional Management System, an embryonic system of competency-based instruction, and a school improvement project using techniques developed by SRI, Incorporated. The Evaluation Branch works with other divisions of the school system, chiefly the Planning and Development Branch, in helping local schools use information from these activities.

Training provided for team members covers three main topical areas: Educational Planning, Group Process, and Budget and Finance. The nine training modules offered in Educational Planning reflect the broad scope of the local school budgeting process:

- Needs Assessment
- Setting Local School Goals
- Prioritizing and Evaluating Possible Programs
- Making the Best Use of Limited Resources
- Soliciting External Funds
- Building Volunteer Support
- Strengthening the Program through Community Resources
- Capitalizing on Parent Interest
- Evaluating Program Outcomes

Although those titles might suggest that we are trying to make expert program planners out of all LSB members, such is not our intent. Indeed, in some communities, we are satisfied if a quorum of a team membership attends most meetings. On the other hand, we do have some communities in which the teams have built on the training content with an enthusiastic and dedicated effort to tailor the program to the needs of their students.

What characterizes the settings in which LSB works best? We have found that the single most important component of a successful LSB effort is the wholehearted commitment of the local school administrator. The principal is most important because he/she alone can weld the staff, community, and school clientele into a cooperative unit. The process simply does not work in settings where principals give only condescending endorsement to collaborative planning.

A second obviously important ingredient of LSB successs is community interest. Our underlying assumption has been that none of our local school communities is without an adequate degree of interest in education to permit the process to work. In a few instances we have yet to validate this assumption. We have reason to believe, however, that distrust on part of some segments of the community is a major reason for non-participation. Also, many residents of the poorer sections of the community seem to doubt their own ability to contribute to such a process.

A question we have often been asked is, "What does a process like LSB cost? That's hard to answer. In one sense, it costs everything because the scope of the planning effort eventually entails every element of the district's budget. In another sense, it costs nothing. The entire process has been effected by realigning resources and setting new priorities. We have spent no more money on LSB than we would have spent without it; remember, our poverty was a key motivator for taking this route in the first place. We believe that the effort to better direct our educational resources to the needs of students at the local school level can only increase our cost-effectiveness.

The Local School Evaluator

The agent serving the evaluation needs of the schools is called the Local School Evaluator (LSE). This is a centrally-based staff member assigned responsibility for helping the local school evaluate its educational program. LSE's have been functioning in Cincinnati for ten years.

The position of LSE was first operationalized within Title I. Approximately half the resources formerly invested in the centralized evaluation of

Title I were diverted to serve local school needs through this position. Lack of available funds has limited the implementation of this function in other school settings. With the advent of LSB, however, the service has been extended.

Each LSE is assigned a group of schools whose respective evaluation needs he/she must serve. Early in the school year, the LSE makes contact with the assigned schools and outlines available services. The functioning of the LSE for the year is determined in large measure by the specific services requested by the local school.

In cases where response to the initial contact is slow, the LSE will attempt to stimulate further interest by offering assistance in addressing specific local school problems. These offers frequently take the form of letting one school know of a service that has proven valuable in another location. The history of local school evaluation in Cincinnati indicates that the services of the LSE are valued as a means of improving instruction.

Ahn, Barta, and Rockwood have described the LSE function as an evaluation track of services supporting the instructional track. Needs assessment supports the selection of content and goals; diagnosis of entry behavior supports the specifications of objectives. Assessment of implementation supports the selection and implementation of strategy, group, time, space and resources; and assessment of end results supports the instruction process itself. (5)

---

5. Ahn, Unhai R., Barta, Maryann B., & Rockwood, Stacy F. "Localizing Evaluation for Public Schools: Concept and Practice." Presented at Annual Meeting of the American Educational Research Association, San Francisco, CA

To give the reader a better understanding of how the LSE assists the local school's planning effort, let me describe six specific services selected as representative of LSE functioning. Three of these serve the process and product of evaluation functions and three the context and input evaluation functions.

The examples will reflect how the LSE interacts with school staff, in addition to the LSB teams. Staff interaction is an important part of the LSE function; its significance may be obscured by the emphasis of our earlier discussion.

Alternative Data. Special procedures for evaluating Cincinnati's alternative programs call for an evaluation team to study each program on a cyclic basis. The LSE is required to furnish the evaluation team with relevant data. In the course of the team visit, the major focus is on process and product evaluation. The visit begins with a discussion of data that the LSE has assembled, and the interpretation of this information by the evaluator is crucial to setting the tone for the visit.

Data forms have been devised for summarizing the data on the alternative programs in accordance with the set of alternative questions prescribed for team responses. Of primary interest here are those questions that generate comments and ratings to be shared only at the local school level. These questions concern curriculum, administration staffing needs, space utilization and concerns of persons involved.

For example, the LSE provides data to answer the question, "Does the status of reading achievement scores suggest a change in overall program emphasis?" Another question related to curriculum is, "Do item data reflect certain skills that need increased emphasis or modified approach?" Related to student concerns, the evaluator compiles data to answer the question, "Do students' survey results suggest specific concerns that the program or staff should address?" In the course of the visit, team members also seek an answer to the question, "Does casual student questioning reflect other matters of important concern to students?"

CIMS Reports. Another service of the LSE related to process and product evaluation is assisting staff in interpreting and applying reports from the Cincinnati Instructional Management System (CIMS). This is a computerized system of monitoring student progress in achieving fundamental reading and arithmetic skills. The teacher using CIMS inventories can more easily keep track of each student's progress in skill mastery. The evaluator can assist the teacher in understanding the computer-generated reports, recognizing common needs of students and planning strategies to serve these needs.

One CIMS report, for example, indicates students' mastery levels on each of a variety of reading skills. These include skills in phonetic analysis, structural analysis, vocabulary, literal comprehension and interpretive comprehension. If several students reading at approximately the same level show a weakness in distinguishing hard and soft g sounds, the teacher can arrange practice in that skill for these students. It is this kind of focus that the evaluator helps to make possible in working with CIMS.

Chapter I School Profile. Because of the availability of resources for Chapter I evaluation, each Chapter I school has a larger amount of data than other schools in the system. Over the last several years, LSE's have worked to help school staffs use this information for instructional improvement. One important part of this approach has been the Chapter I school profile.

Essentially, this profile helps the staff of an individual school evaluate certain process and product outcomes in the light of comparison with other Chapter I schools. This has been a recommended approach among these inner-city school units, which formerly suffered from comparison with schools in more affluent communities.

The school profile includes information on student achievement as measured by standardized tests. Information on student attendance; teacher, parent, and student attitudes; data on parent involvement -- all of these are included in profiling the Chapter I school against the background of other schools in the program.

Item Report. One of the most frequently requested services of the LSE is training in the use of standardized test item analysis. The LSE helps teachers use the results of specific items to diagnose student and group needs for instruction or remediation.

Several approaches have been devised to make this process easier for teachers. Among these are the pupil item report worksheet and the group report worksheet. The pupil worksheet allows the teacher to list names of students whose item results suggest a need for more practice time, further diagnosis or other reinforcement. The group worksheet permits the comparison of class unit or school unit item data with similar data for the school system.

Survey Information. Our earlier discussion of goal setting mentioned the work of the LSE in helping local teams use information collected through the annual surveys. The surveys of administrators, teachers, parents, and students provide a large volume of attitudinal information which the LSE can help the local team understand and apply.

The student survey, for example contains a number of items related to self-concept. By comparing the results of these items from one year to the next and with results from other schools, the team can gain useful insight into specific needs for program emphasis.

School Information System. The survey results on each school are stored in the district's School Information System. This data bank includes several hundred variables on each school unit. A number of these, e.g., drop-out rates by cause, socioeconomic variables, etc., are very useful in context and input evaluation.

Each year the school receives a comprehensive report of its data as contained in the School Information System. The sheer volume of this report makes it important that someone be available on a consulting basis to help staff and team interpret and apply the information.

A Look Ahead

What we have seen happen in our attempt to support community efforts to improve instruction has been encouraging. We see considerable basis for hope in this approach. As the educational dollar continues to become more scarce, the communities to whom the schools belong are joining in the concurrent missions of providing adequate resources and putting what we have to the best possible use.

Advisory committees are not new, of course. Many districts have used them, some under requirement of state law, and they have had varying degrees of success in realizing educational benefit from the work of these groups. We see our emphasis and procedure as somewhat different from those of most other districts. And the support evaluation provides to these teams is a key element in this difference.

Perhaps more than in any other district, communication of evaluative data to local clientele is given top priority in Cincinnati. This emphasis has led to long-range planning in which research, testing, and evaluation functions are the means of obtaining information to be communicated to those interested in the schools.

The five-year plan for the school system includes a major section on communications. Four goals are set down for this part of the plan: 1) improve the school system's ability to hear the voice of the public it serves; 2) obtain better information to serve the need of staff, clients and community; 3) improve the quality of information processing and reporting; and 4) assist staff, clients, and community in interpreting and using information for decision making.

Our attempt to support community improvement of instruction has not been free of problems. Stirring up apathetic communities, gaining commitment of sufficient time, internal coordination -- these and other difficulties have kept us aware that we are still in a real world where panaceas seldom exist. We deal with these problems as they appear, using the resources described in this exposition. The success we have had leads us to believe that some of what we are doing may be helpful to managers in other districts.

A superintendent who shares Cincinnati's pressing financial dilemma may be able to apply our strategies for gaining additional community support. Another manager confronted with non-productive advisory committees could perhaps use some of Cincinnati's techniques for involving them more integrally in educational planning.

Whatever the application, I hope my attempt to describe the LSB process, giving the answers to many of the questions we are asked about it, will be of some utility. I hope that the reader can benefit from what we have learned about the key role of the local school administrator in applying such a process. I hope that our experiences in trust-building, working toward consensus, and applying evaluative findings to program choices will be helpful.

I hope, too, that the reader can profit from some of the mistakes we have made. It is important to avoid the pitfall of trying to accomplish too much change too rapidly. Especially with a process such as this, it is crucial to allow time enough for the participants to work through many details of operation. Involving all parties affected is a constant challenge. And, most importantly, the selfish interests and protected domains of individuals must never be allowed to interfere with what is really best for the students we serve.

CHAPTER 5

School District Analyses

in which the impediments to developing such systems are discussed along with the conditions propelling districts in this direction . . .

CHAPTER 5

SCHOOL DISTRICT ANALYSES

Introduction

In this chapter we take a broad conceptual view of why school districts do not naturally link their testing and evaluation with their instructional programs, even though this may seem a logical or important linkage to develop. We include this section for several reasons: one, we believe that you cannot improve or change something unless you understand it; two, we want to dispel some common misconceptions about how school districts operate and change; three, we want to stress the importance of school districts developing their own approach to linking testing, evaluation and instruction rather than adopting ready-made programs.

First, let us consider the steps commonly taken by school districts that are found with low or declining test scores. Again, we will refer back to the scenario in doing so.

Why Districts Have Trouble Dealing With Test Scores

In developing this section, we will start off by referring back to the scenario in Chapter 1. While the scenario focuses on the decline of test scores, it also contains several elements that are typically found in school districts and that complicate attempts to reach a solution to this test score dilemma. For example, consider the following:

1. Lack of consensus over the problem's cause. A district's declining test scores can be caused by one or several factors: poor instruction, inadequate time on instruction,

limited match between what is taught and what the test measures, statistical variations such as standard error, lack of pupil knowledge of how to take tests, shifting pupil population. Each of these variable causes suggests a different solution or strategy.

2. These declining test scores are not the only problem the district faces. The electrical power failure in high school demands immediate attention. The district is experiencing declining enrollments which will likely force the district to initiate the difficult task of closing schools and possibly laying off staff. Both of these processes, closing schools and laying off staff, are generally politically and organizationally difficult problems that will likely consume much time and effort. Of course, districts often face other problems, e.g., desegregation orders, drug problems, contract negotiations, that also chip away at the amount of time and energy that testing problems can receive.

3. The district does not have the luxury, or freedom, to present the problem to the public in its own terms. In this case, the newspapers presented the story and likely influenced considerably the way the the public interpreted the results. Newspaper reporters generally are not educators

and sometimes, but by no means always, their interpretation of educational events can be limited or reductionist in character, e.g., the teachers must not be teaching the basic skills. While lack of teacher attention to basic skills may not really be the problem, the district must, nonetheless, deal with the public's perception that this is indeed the problem.

4. School management cannot take unilateral steps to solve the problem. In this scenario, the school board, or a representative public body, and the teachers' union, which represents the school teachers, will likely have a major say in whatever is decided. The board, for example, if it is concerned about re-election, may insist that the district's solution to the problem be in response to the public's perception of the problem. Thus, they may insist on more classroom time on basic skills. The teachers' union, on the other hand, may insist that teachers be paid for any out-of-school inservice training sessions. If the district is facing budgetary problems, this may effectively remove inservice training as a viable strategy for solving the problem.

5. Lack of consensus on the solution to the problem. If one can not define the cause of the immediate problem it follows that it will be impossible to select an appropriate solution to the problem. But even if one can reasonably define the problem, there is limited consensus about what

> to do about it. For example, if the problem is defined as inadequate time on instruction in basic skills, does that mean that all pupils should increase their time in basic skills instruction or just a few? The instructional time in schools is finite; if instruction in basic skills is increased, what other subjects will "lose" instructional time? How will that dilemma be resolved?

Given all these characteristics of school districts, is there nothing that can be done? Must school districts continually be buffeted back and forth on the basis of test scores? Must they forever be placed in a dilemma where they cannot define the problem, have limited choice over solutions, and have no realistic way of determining whether or not selected solutions are achieving the intended results?

Some districts, in response to declining test scores, take specific steps, such as: teaching the pupils how better to take tests, or increasing the required classroom time on certain subjects such as reading or math, or initiating yet another series of inservice training workshops for teachers. We would argue, these isolated steps generally do not make much of a difference. These steps do not usually improve test scores very much and the dilemmas, and sometimes risks, remain. In the mean time, teachers begin to lose their feeling of efficacy, the public lessens its support for the schools and the pupils are judged as not receiving an adequate education.

We think not. While the problem is complex, it must be solved if districts are going to make progress in providing a good educational program while at the same time improving public confidence and satisfying staff concerns. But it is not easy. School districts, faced with such a problem as that described in the first chapter's scenario, are confronted with a formidable problem. As the superintendent and board members and staff members in our scenario meet to consider the problem, they must address several questions.

How can we describe the problem? Where do we start? Who should be involved? How will we know if we are making progress or have solved the problem? The remainder of this chapter deals with some basic assumptions that we feel should be kept in mind when this or any school district approaches this problem.

Some Assumptions

As the administration, board members, and staff members in our scenario school district sit down to begin to ponder their dilemma, we would recommend that they keep in mind some assumptions that should form the basis for their discussions. We will list the assumptions here and then subsequently discuss each in more detail immediately thereafter.

> Assumption 1. The problem cannot be solved unless it is first analyzed and understood.
>
> Assumption 2. The problem is a school district systems problem and not the fault of any one individual, groups of individuals or single process.
>
> Assumption 3. A strategy for linking together the district's testing and evaluation and instructional program can be an effective strategy for solving the challenge posed by declining test scores.
>
> Assumption 4. A system for linking testing and evaluation and instruction must be situation-specific and accommodate various external and internal factors.
>
> Assumption 5. Developing and implementing a subsystem for linking together the district's testing, evaluation and instructional programs (t/e/i linking subsystem) can be thought of as an educational innovation.

Assumption 1. The problem cannot be solved unless it is first analyzed and understood.

This seems like a self-evident statement, yet we have observed and studied district after district that has adopted a particular strategy or taken certain steps without having taken this first problem-identification step. In order to satisfy public demands for action districts often choose a single strategy such as mandating that instructional time on selected subjects be implemented, or changing the district's reading or math or social studies program, or requiring teachers to attend mandatory inservice programs, or changing the norm-referenced tests the district is using. This list of selected isolated activities could go on and on. The district's basis for these seemingly isolated strategies might include such factors as an influential book salesman or a local university consultant or some unexamined belief system of a prominent school board member or administrator.

In making these statements we don't mean to be hyper-critical of school districts who act in this manner. As we pointed out earlier, and as was illustrated in the scenario, schools districts have many problems pressing on them for solution. Often there simply isn't time to give this problem the sustained attention it needs, especially if the district is unsure of, or lacks the technical capacity to identify and deal with this problem. Yet, the problem will not go away and its potential ramifications can be serious indeed.

Let's look again at the scenario district for ways in which the district might begin looking at the problem. The most immediate question is, "What is the problem?" or perhaps more accurately, "What are the problems?" Many districts often assume that the lowering test scores automatically mean that the students have learned less or that the quality of instruction has declined. Of course, it may not be an instructional or learning problem at all. It may be that there is a whole relationship between what the test measures and what is taught in the classrooms. Either the instruction should change or a different test might be selected. Or districts make comparisons between test scores from one year to years past. If there is a decline, they assume that the schools are not performing as well as in years past. But is the school population the same? For example, there may have been a significant influx of new students whose previous educational background is substantially different from the district students who were tested earlier. What the test results might indicate is that the district test does not measure what new, or relatively new students learned elsewhere. Or, a district might conclude that the instructional program simply doesn't give adequate time to certain subjects. But often times, districts have no basis for knowing how much time is presently being spent in classrooms on various subjects. Or a district might conclude that a present instructional program isn't effective, but they have little knowledge about whether or not the program has been implemented by the teachers as it was designed.

In a sense, what starts out to be an essentially straightfoward problem quickly dissolves when various alternative problem statements are considered. In a sense, the "testing problem" becomes an evaluation problem. The situation cries out for a thorough systematic assessment. The dilemma will not likely be dealt with effectively until such an analytical process is begun. Please note, we are not suggesting that such an analysis is a single, massive effort to analyze the problem. Rather we are suggesting an analytical frame of mind that will result in long-term, continuing inquiry into the problem, and the results of various strategical implementations to solve the problem.

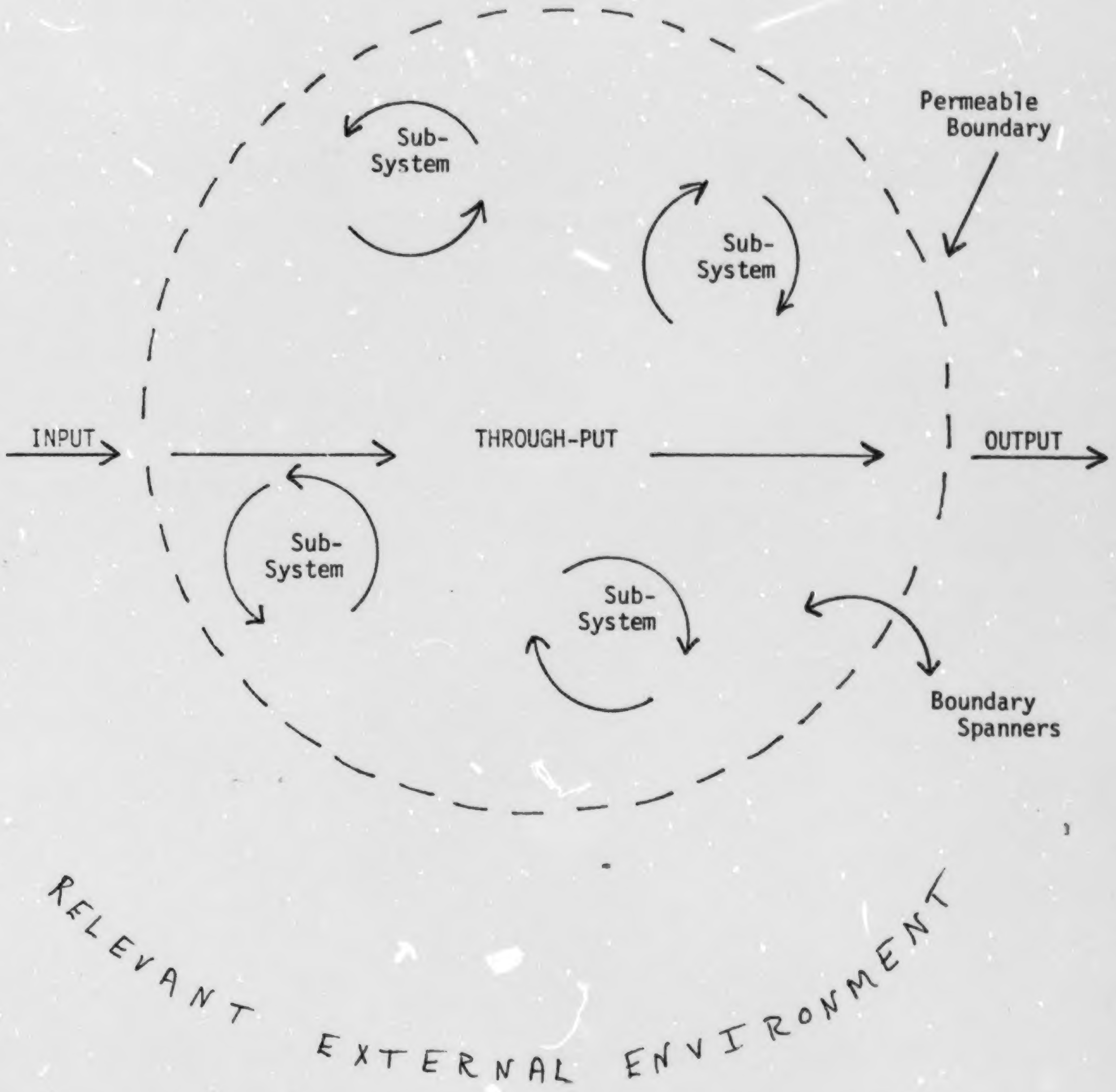
Assumption 2. The problem is a school district systems dilemma, and not the fault of any one individual, groups of individuals, or single person.

School districts, even small ones, are complex organizations. It is difficult, often, to conceptualize the problem because we limit our thinking to the district organization itself, with its school board, school sites, and administrative and teaching staffs. Often, we focus our attention on some subsystem, such as the central administration or school sites when conceptualizing the problem. But, of course, school districts do not exist in isolation from their environments. Nor do the various subsystems exist independently of each other.

But how does one conceptualize such a complex entity? We find the open systems model of organizations to be useful in this regard.

School Districts As Open Systems. Open systems is a widely-used and relatively simple conceptual model. It is used in many fields, such as biology, to illustrate the functioning of an organism in its environment. One of its major features is that it accounts for the external environment's influence on the unit, or organization being studied. By using the open systems model, one is forced to consider the dynamic interaction between an organization and its environment.

Viewed graphically, a school district as an open system can be depicted as follows:


Permeable Boundary
Sub-System
Sub-System
INPUT
THROUGH-PUT
OUTPUT
Sub-System
Sub-System
Boundary Spanners
RELEVANT EXTERNAL ENVIRONMENT

Relevant External Environment. This includes those agencies, individuals, groups of individuals and other organizations that can influence an organization's functioning. In the Chapter 1 scenario, this was represented in several ways: parents calling the schools, the newspaper article, community persons. Of course, the list of external environmental influences is much larger, and includes such groups and individuals as: the federal and state courts; federal, state, and local governments; taxpayer associations; radio and television; local community and business leaders; post-secondary schools to which the district's population matriculate and local private schools. From time to time, these groups and individuals can exert considerable pressure on the school district to confrom to its or their wishes. Many factors influence the district's response including the external groups, political power, legal mandates, and the saliency or popularity of the issue. What is more, the district's response to these demands, and, some have argued, the district's internal organizational structure, must respond to the make-up and complexity of the external environment if the organization is going to deal effectively with its problems.

Permeable Boundary. All organizations have a boundary around them that separates the organization from its external environment. It is, however, almost impossible to determine exactly where that boundary is. In the case of the schools, for example, the governing body of the district (that is, school boards) are agents of the state. Does it include the parents who in many instances have final authority over the child's education, but limited authority in other instances? This issue is essentially unresolvable, depending as it does on the problem or action being considered and how one defines influence.

This boundary is permeable, that is, virtually no organization is completely independent of its external environment. External influences can penetrate the boundary and effect its operations and decisions. But the permeability of the boundary varies considerably from organization to organization. In the case of public school districts, the boundary is quite easily bridged. That is, many of the school's actions on such vital matters as budget allocations, whether or not to close schools, the school's test scores, are easily accessible to the external environment. What is more, decisions by external agencies such as federal courts or state legislatures and local business, for example, that may decide to open a new factory, often bridge the school district's permeable boundary and require that the organization vary its processes or decisions.

In our scenario, this boundary penetration is clearly illustrated by the way in which the test score data were obtained by the press and subsequently by the effect the newspaper article had on the external environment and its subsequently penetrating the organization to cause it to consider the appropriate action to be taken. This boundary permeability has both advantages and disadvantages. On the plus side, it forces an organization to be responsive to items and issues that the external environment considers important. Organizations cannot ignore the external environment and hope to remain current with the needs of the constituents they serve. On the negative side, the organization can be easily buffeted by changing public attitudes that may be based on incomplete information or on faulty analyses, or on developments elsewhere that may be inappropriate to the school district's situation. The result of this is that school districts find it difficult to conduct long-term planning; a plan that is developed and implemented based on the public's view or wishes at a given time may have to be altered or discarded before the plan is fully implemented because some influential agency or group in the relevant external environment has changed its goals or priorities.

Internal Subsystems. All complex organizations, such as school districts, are composed of subsystems. This factor is sometimes forgotten when we think and speak of organizational actions and decisions. Then, we say a school district did this or that, or performs in certain ways. We would be more accurate to say that various subsystems and individuals have decided something or are performing in a certain way. Thus, a school district can have a curriculum that is adopted by the school board (one subsystem), articulated by the curriculum department (another subsystem), and implemented by the teachers in the school sites (another subsystem). And, as will be elaborated in a subsequent section, the articulation and coordination among these various subsystems is often less than perfect. Thus, one can talk about the district's curriculum, but general statements about it often ignore or gloss over the complexity that is inherent in such a general statement.

A complex organization then consists, internally, of a whole series of subsystems that vary considerably in the degree to which they are working toward common goals and the level of coordination and cooperation among them.

Input, Throughput and Output. This organization of operating subsystems, which exists in a dynamic relationship with its relevant environment, then, takes in resources (input) in the form of money, materials, personnel, clients, etc. which it parcels out into organizational subsystems, in the form of such agencies as school sites, personnel offices, curriculum units, budget offices, custodial and support staff, etc., to produce an output, such as institutional programs, educated pupils, employed citizens (both graduates and citizens who work in the district), community entertainment (such as athletes and musical performers), and babysitting (that is, pupils are engaged and supervised during the school day).

While there are commonalities from district to district on these inputs, throughputs and outputs, significant differences exist among districts that make generalizations hazardous. That is, schools differ considerably on such important items as: level and source of financial support, mix and variety of pupils in racial and socioeconomic terms, district organization of the subsystem, coordination among the subsystem, definition of a successful graduate and organizational goals.

Summary. We believe this open systems model of school districts provides a comprehensive way to look at how a school district functions and the various initial factors that must be considered if the organization is to deal effectively with the problems and challenges it faces. In the case of our Chapter 1 scenario, one begins to see the interplay among the various parts of an open system. The district is seen as a complex of internal subsystems that are being effected dramatically by the dynamic interchange of the organization with its relevant environment across its permeable boundary. The various subunits are being called together to forge a response to the challenge -- the declining test scores.

One final comment. It is important to note that this is an open systems model, and not a theory. That is, it provides a way to conceptualize how complex organizations function in their environment. It provides, we believe, an effective way to view the organization's functioning but it does not in any way predict how an organization does or should respond to external challenges. Organizations indeed respond in many ways to their relevant environment. The exact relationship between their response and their organizational success is at best speculative. We believe that one way the district can respond to the declining test score challenge is to consider linking its testing and evaluation program to its instructional program. It is to this process that we turn next.

Assumption 3. A strategy for linking together the district's testing and evaluation and instructional program can be an effective strategy for solving the challenge posed by declining test scores.

The wording of this assumption is important. Note that we have said that a t/e/i linkage strategy can be an effective strategy for solving the challenge posed by declining test scores.

We are not suggesting that this t/e/i linkage is the only way a district can meet this challenge. Some districts might, for example, simply not give norm-referenced tests and release the results to the public. Others may decide that strategies that could be considered "humanistic" would be more appropriate. We have no arguments with those who would select such approaches. Our advocacy of this t/e/i linkage strategy is based on what we perceive as the public's belief in the importance of norm-referenced tests, or criterion-referenced tests, or competency tests or some combination of these. (For a more detailed discussion of tests and what they measure, see pp. ___.) For good or bad, testing has become accepted practice in American schooling. Given the pervasiveness of this practice, it makes sense, we feel, for districts to maximize the potential usefulness of those tests in informing decisions about the instructional program.

Program Evaluation, though a more recent phenomenon, is also a common practice in many school districts. Largely in response to federal and state evaluation mandates that accompanied external funds for special programs (such as ESEA title I, ESEA Title IV and PL 94-142), school districts have either developed a capacity for in-district program evaluation research or they have developed a continuing relationship with an external evaluation agency that can provide systematic evaluations. Many districts now have the capacity to perform evaluations and, with some modifications, these evaluations can serve to inform local school district instructional decision making.

Also note in our statement of this third assumption that we refer to this as an "effective strategy for solving the challenge posed by declining test scores." This wording intentionally refers to the declining test scores as a challenge. We do not take the position that the declining test scores are themselves the problem. Instead we view them as symbols or manifestations of the instructional program. Our interest is in the information presented by test scores and through evaluation to improve instruction. A district could have an excellent instructional program and poor test scores. Similarly, a school district could improve its test scores, through various manipulations of students, teachers and statistics, and make no instructional improvement. The key concept here is that the testing and evaluation activities and capacity can be linked together to inform instructional decision making, and thereby improve instruction.

Our basis for suggesting this t/e/i linking strategy is based on developments on a national scale and at the local level in many districts in testing and evaluation.

In testing we have made considerable progress in improving the quality of norm-referenced tests and in understanding their strengths and weaknesses in providing information about instructional programs. In short, there is no need for districts to blindly react to norm-referenced test scores as though they are the ultimate measure of a school district's program. The fact that some districts or communities continue to use norm-referenced test results in that way suggests that the present sophistication regarding those tests' strengths and weaknesses is not uniformly known or adhered to. Norm-referenced tests can provide important data but only if the relationship between the norm-referenced tests' content and the school district's instructional program is known and understood.

A parallel testing movement has developed in the design and use of criterion-referenced tests. These tests can be more specifically related to a district's instructional program and accordingly they can provide almost immediate feedback regarding the district's program. But their potential use for informing instruction is limited to the degree to which they are linked to the instructional program. As with norm-referenced tests, there are certain strengths and weaknesses associated with criterion-referenced tests. But if properly designed and linked to the instructional program, they can provide valuable data.

Still a third movement is the competency testing movement. While somewhat more limited in scope than the norm-referenced and criterion-referenced tests, these too can provide valuable information about the instructional program's ability to provide a basic survival education. Data collected from these tests can answer important questions about the school district's instructional program and the students' progress in achieving basic survival skills. No one will argue that this proficiency testing program should be the ultimate goal for the district, but it does provide a base line upon which the district can build. And, data from these tests can provide valuable information for district instructional decision making.

We make no argument for one testing system over another. We only remind the reader that each test has its strengths and weaknesses; accordingly, each provides its own kind of data (see pp. ___). And these data can form an instructional management information system but only if they are adequately linked to the district's instructional program.

Program Evaluation is a more recent development in education. The main emphasis for its development came from requirements that accompanied federally, and subsequently state funded programs for systematic evaluation of such projects. The purpose of these evaluations was to provide federal and state officials data upon which to make decisions about whether or not to continue program funding.

These requirements started a sizable movement in education. A number of scholars, many of them trained in educational psychology and psychometrics, began developing methodologies and procedures for program evaluation. Educational evaluation began to develop as an identifiable field of study replete with a Center for the Study of Evaluation, its own division in the American Educational Research Association, graduate programs in major universities, and both research and practitioner journals.

Paralleling, and indeed a major stimulant for this development of educational evaluation as a field, was the development of school districts' capacities to conduct educational evaluations. In the beginning of the movement many school districts handled all the evaluation requirements by utilizing external consultants or by adding this task on to the duties of those who handled the district's testing program. This procedure continues in many districts, especially smaller districts, today. Many other districts began to consolidate all the evaluation funds into a separate unit which was given a variety of names which included some mix of terms such as research, evaluation and testing.

The result of this movement was to provide districts with increasingly sophisticated capacity to conduct program evaluations. But recent research has revealed that much of the work of these units, that is, their evaluation reports, has been directed, understandably, to external funding agencies. The evaluation reports have been designed to answer the external funding agencies' questions; they have not been developed to meet the local needs. Also, the evaluation community has come to realize that in many instances the evaluations were not being utilized as intended by decision-makers.

There were many reasons for this, such as a misunderstanding of evaluation's role in decision making and the evaluators' adherence to standard research formats rather than developing evaluations that were responsive to the decision makers' needs or pressing questions. In recent years, the evaluation community has gone through reappraisal of its purposes and methods in an attempt to increase the utility of evaluations for decision makers. Considerable progress has been made in that regard.

Thus, presently, many school districts have developed program evaluation capacities consisting either of groups of evaluation consultants or an internal research, evaluation and testing unit. And, this evaluation capacity can provide much needed data on the school district's instructional program if it is appropriately linked to it. That is, this capacity has habitually been used to ratify external funding agencies -- the challenge, and promise of school districts now is to forge a linkage between their evaluation procedures and local school district instructional decision needs.

In summary, school districts are under considerable pressure to improve their instructional programs. Yet, in many -- perhaps most -- districts, instructional decision making is done quite unsystematically, often in response to some set of scores from an annual administration of some set of norm-referenced tests. Yet, these tests have limited capacity for making well-informed instructional program decisions. This is especially true if the relationship between the norm-referenced test and the district's instructional program is unkown. What is more, most school districts, given their organizational pattern with teachers functioning quite autonomously behind classroom doors, have surprisingly little data available about what their instructional program really is and how it is implemented. But developments in testing, such as refinements of norm-referenced tests, the design of criterion-referenced tests and competency tests, can provide districts with useful tools with which to assess the outcome of their instructional program. Similarly, advances in the evaluation field make it possible for districts to begin to obtain systematic and comprehensive analyses, both formative and summative, of the ongoing districts' instructional program. Together, this testing and evaluation capacity, if linked with the instructional program, can provide a district with an instructional managment information system. The fact that this has not occurred in many districts is probably due to several factors including characteristics of school districts as complex organizations, a topic to which we turn next.

Assumption 4. A system for linking testing, evaluation and instruction must be situation-specific and accommodate various external and internal variables.

We suspect, and indeed our research confirms, that few districts have linked their testing and evaluation activities with their instructional programs. There are probably a number of reasons for this, such as not realizing the potential or the district's capacity to develop such a linkage or simply the district's having decided on another strategy for instructional improvement. We suspect, however, that the characteristics of school districts in general are such that they make the linkage of these three processes, testing, evaluation and instruction, difficult to accomplish. We suggest that the following four characteristics of school districts mitigate against school districts moving naturally or easily toward developing this t/e/i linkage: institutionalization, goal differences, permeable boundaries, and loose coupling.

Institutionalization. This term, derived from the sociology of organizations, suggests that some organizations are valued by the public as institutions and given special status and recognition. Generally they receive public funding and a virtual monopoly for their activities. They do not have to compete, in the typical sense, for survival and they are under less pressure than most other organizations to justify their existence. Examples of such organizations would be national parks, public libraries and public schools. This special status, of course, confers to these organizations certain advantages in that they do not have to continually fight for survival although, to be sure, the level of their support is subject to political trends. They are by no means trouble-free.

But this institutionalized characteristic has its disadvantages as well. That is, the organization can survive without looking critically at the way that it operates and the degree to which it is successful in accomplishing its goals. A common outcome of this institutionalized characteristic is that institutionalized organizations often deal in symbols rather than in reality. Schools and school districts, when asked if they are of high quality, will usually respond by referring to such factors as: the fact that all the teachers are credentialled, or that an admirably high percent of the teachers have Masters degree, or that the libraries have so many volumes, or that the curriculum is clearly described. Seldom are questions raised as to whether or not credentialled teachers or teachers with higher degrees are more competent than those without such symbols, nor do they really see if pupils are reading and understanding all those books or if the curriculum is actually being taught in classrooms. In other words, schools, as institutionalized organizations, tend to deal in symbols -- they pay surprisingly little attention to their instructional core. Teachers, as a group, receive little supervision over their instruction.

What is more, some have argued, it is not in the schools' best interests to actually look very clearly at their instructional core if that core is weak. And it is widely recognized, that the actual technical core of teaching is quite weak. For example, educators have few really carefully researched and developed methods to deal with the challenge arising from an influx of many non-English speaking pupils. Indeed, there are many different viewpoints both about the advisability and techniques of bilingual education. In saying this, we do not suggest that educators are uncaring or that a little thought and research has not taken place with regard to this issue. Instead, we suggest that bilingual education, as an example, is a value-laden issue and a very complex instructional problem. Whatever the reasons, the fact is that schools have a very weak technical core. If schools look very critically at that core, and expose it to public scrutiny, the schools may begin to lose their institutionalized status. Some things are best left unexamined! In summary, this institutionalized characteristic of the schools lessens the schools' incentive to link testing, evaluation and instruction.

Goal Differences. The public schools have long suffered from lack of agreement over the goals the schools should pursue. Readers with a sense of educational history can remember the movements from humanistic education to vocational education to basic skills. What is more, the goals of one segment of society will often differ from those of other social segments. While there is general agreement on, say, the basic skills, there is considerable disagreement over priorities and over the importance and the best strategies for pursuing such other schooling functions as racial integration or developing critical thinking skills. Thus, schools are constantly subject to a variety of external pressures to conform to the decision of one group or another. Given an open system orientation, it is important that they attend to their clients' decision, but, as was pointed out earlier, these constantly shifting priorities make it difficult for districts to mount sustained, long term and carefully researched efforts to implement a given goal. Thus, school districts tend to sort of muddle through, bouncing back and forth between various changing demands that are placed on them. Institutions in this kind of situation tend to disassert the need for a carefully designed management information and planning system -- it is simply not worth the time and limited reserves it consumes.

Permeable Boundaries. This feature has been described in some detail in an earlier section of this chapter so we will not elaborate further in describing it. We should repeat, however, that this permeability makes it difficult for school districts to protect themselves from external activities that both change the basic direction of the institution or direct a good deal of the district's attention and reserves away from the central instructional program. Many school districts, for example, have had to deal with demands placed on them from the external environment for which they have been unprepared and which have demanded a good deal of their time. Examples are school programs to handle: desegregation, educating handicapped pupils, large immigration of Latino or Asian pupils. We do not question the merits of such programs, only that they are so large in scope and complex in execution that they demand an enormous amount of time and resources and they force the schools to spread their planning and development capacities very thin. Districts so occupied with external demands are not likely to have the energy and resources to attend to their central instructional core, especially if the institutionalized character of the organization reduces the immediate need for them to do so.

Loose Coupling. A currently popular term, used to describe the schools, loose coupling refers to the characteristic of some organizations to have loose couplings, or connections, among the organization's various parts or subcenters. Assume that an order is given, by an organization's management, that a certain procedure or change is to be implemented. In a tightly coupled organization one will expect that the order or change will be carried out within a reasonable period of time. In a loosely coupled organization, the order or change may never be carried out or done so in a sporadic or idiosyncratic manner.

No organizations are completely tightly or loosely coupled. In schools, for example, if the school board orders that the beginning and ending hours for the district schools be changed for the next school year, one can be almost certain that the order will be carried out. In that instance, the organization is tightly coupled. In the case of the instructional program, however, if the school board orders that teachers institute a particular instructional approach or base their instruction on the results of certain tests, one cannot be at all sure that after a reasonable time the order will be carried out. The school's instructional program is often loosely coupled from the organizational management system.

There are, of course, reasons for this, the main one being that the organization's main technical core unit, the classroom, is operated quite independently by teachers. District managers have neither the resources, inclination nor agreed upon technical superiority to substitute their judgment for the teachers' over instructional matters. What is more, the various district subunits, especially in large districts, tend to operate in considerable isolation; they are themselves loosely coupled from one another. This loose coupling makes it unlikely that the testing, evaluation and instructional processes can be easily or naturally linked.

Assumption 5. Developing and implementing a subsystem for linking a district's testing and evaluation together with instructional program (t/e/i) can be thought of as an educational innovation.

As was pointed out in previous sections, it is not natural for school districts to link testing, evaluation and instruction together. There are many institutional, organizational and historical factors that mitigate against such linkage. If such a linkage is going to be forged it will take some positive, direct and sustained planning and action. For some organizations, e.g., a manufacturing plant or an engineering firm, the use of an information management system that consists of the regular monitoring of critical technical core processes to determine whether they are being completed as intended and whether or not they are having the intended effect may seem to be a standard operating procedure. Presently, in public school districts, however, such a process can be considered an innovation or change.

During the past twenty years or so a great deal of research has been completed on the educational change process. While the resultant research literature is too extensive to review here, there are a few generalizations that seem to emerge from that literature that would apply to districts wanting to develop a t/e/i linking subsystem. The few generalizations that we feel apply are: the importance of ownership, the incremental nature of comprehensive change, the process of mutual adaptation and the role of incentives.

The Importance of Ownership. As the name implies, ownership simply means that those who will be affected by an innovation have some sense that the innovation belongs to them. It conveys a feeling that those affected have at least been consulted about the innovation, or better still, that they have played an important role in determining why, how and when the innovation will be introduced, implemented and adopted. This ownership can be developed in a number of ways, by having those who will be affected, or some group representing them: be involved in developing the innovation; be participants in deliberations and discussions on whether or not to initially implement the innovation; be involved in judging whether or not the innovation is working as intended; and being involved in final adoption decisions.

Consider the differences between a school district that decides at the top administrative level to conduct, as an analytical step in building a t/e/i linking subsystem, an evaluation study to see if teachers are uniformly spending enough time on mathematics and if they are using what the district considers "appropriate" teaching methods. Assume further that the decision is made quite unilaterally and the teachers are simply informed that this is an important study and that the data will not be used for personnel purposes. One can pretty well predict what will happen. The teachers will be suspicious of the district's motives, they will resent the intrusion into their classroom, the study will become the target of teacher resistance, and this will likely minimize the accuracy and therefore the usefulness of the data gathered.

Suppose instead, that teachers from the beginning were told that the district wants to work with them in improving instruction and that the responsibility for whatever inadequacies there might be in the instructional program belong to everyone including teachers, principals, and central administration. Suppose further, that it is proposed that a joint committee be formed that consists of teachers, principals and relevant central administrators to explore the possibilities of conducting a study to determine current practice. And, if such a study were deemed valuable, that the study's design, implementation and analysis would be directed, though not necessarily conducted, by this joint steering committee. Problems would still exist but one can predict that because the teachers, who are the ones to be studied, are involved there will likely be less resistance, the study will be more realistic in its design, execution and analysis, and therefore more useful to everyone. One risk of this approach is that the teachers may decide that the study should not be conducted because of reasons they consider important. If such is the case then the study should be either redesigned so as to meet the teachers' concerns or forgotten. It is difficult to over-estimate the importance of ownership in designing and implementing innovations.

The Incremental Nature of Comprehensive Change. For many years, it was taken as an act of faith to describe the innovation process as a linear, logical process subject to standard planning methods. Innovators were urged to state clearly their objectives, determine the resources needed, develop a detailed plan for developing and implementing the innovation, and determine the appropriate evaluation methods and ways of feeding evaluation back so that the innovation could be properly adjusted.

While this standard planning procedure makes sense logically and it looks good on graduate examination papers, it bears little relationship to what actually goes on when comprehensive innovations are introduced into complex organizations. There are likely many reasons for this, e.g., many complex interactions are difficult to anticipate, not enough is known about the situation, or the innovation is not sufficiently developed so as to lend itself to such precise planning. We suggest, however, that perhaps the paramount reason is that educational innovations are being introduced into a human system that simply does not respond to such planning. Instead of being logical, linear and predictable, most innovations, when looked at in retrospect, are implemented on a non-rational, non-linear and incremental basis. And the patterns of implementation vary almost as much as the variety of innovations themselves. This isn't to suggest that planning is to be completely avoided. Indeed, one should proceed with an innovation with a plan in mind. However, one should avoid being too wedded to a particular strategy and one should be prepared, indeed expect, that the original plan will likely be changed considerably from time to time before the innovation is fully implemented.

Mutual Adaptation. Related to the incremental planning concept described above, is the concept of mutual adaptation. Simply stated, this suggests that the chances of an innovation being implemented improve considerably if allowances are made for the innovation and the setting to mutually adapt to each other. That is, the innovation of, for example, a new reading program, will likely undergo some adaptation as it is implemented in various sites. It is unrealistic to believe that it can or should be implemented in exactly the same way in different settings. Similarly, it is likely that the school settings will have to adjust themselves, though in different ways, in order to accomodate the innovation into their settings. These adjustments may be in the form of changing personnel or schedules or administrative procedures. It is not likely that any significant innovation can be introduced without some organizational adjustment.

The Role of Incentives. For every innovation there are those individuals and groups who expect to benefit from the innovation and those who will either receive no benefit or will be disadvantaged by the innovation. For those advantaged, there are said to be incentives. For those disadvantaged, the innovation carries disincentives. It is not unlikely for those who are disadvantaged to offer resistance to the innovation. Nor is it unreasonable for them to be less than enthusiastic about an innovation, especially if they perceive the innovation as using them to benefit someone else.

The implication of this is that one should make every effort to design and implement innovations in such a way that everyone benefits, and that the benefits are clearly spelled out to those who are participating. In our own work over the years we have come to view this as a cardinal rule of successful innovation. People will more likely cooperate with those innovations that help them solve a problem that they have. Few people can resist innovations that solve real, pressing problems. Few people will assist innovations that promise no personal payoff.

In summary, those considering the development and implementation of a t/e/i linking subsystem should consider this as being so atypical of most school district functioning as to regard it as an innovation. As such the innovation should be planned and implemented in such a way that those who will be affected by it will feel ownership for the plan, the plan will be sufficiently flexible and sequenced that it can be changed to meet changing local conditions and circumstances, the innovation will easily allow for mutual adaptation and the innovation will be so designed and explained that all who will be involved in the plan will see in it ways to solve real and pressing problems.

CHAPTER 6

School District Possibilities
in which district prsonnel, who are enroute to their own instructional renewal, examine their own situations in terms of impediments as well as facilitating conditions . . .

to be submitted 12/01/82

CHAPTER 6

School District Possibilities

in which district personnel, who are enroute to their own instructional renewal, examine their own situations in terms of impediments as well as facilitating conditions . . .

Chapter 6

School District Possibilities

Sections

1. Improving instruction through the management of testing and evaluation activities: an action planning perspective
    - a) Introduction
    - b) Definition of action planning
    - c) Attitudes on which district action planning rest
    - d) Conditions impeding or facilitating action planning

2. Alternative action plans: discussion and outlines
    - a) Test-Focused
    - b) Curriculum/Instruction-Focused
    - c) School Planning-Focused
    - d) District Management-Focused

3. Tips and Techniques:
    - a) Tests: Criterion-referenced; Norm-referenced
    - b) Evaluation
    - c) Curriculum and instruction
    - d) School planning and assessment
    - e) Management

4. Initiating action planning: Discussion and Worksheets

Section 1

Improving Instruction Through the Management of Testing and Evaluation Activities: An Action Planning Perspective

Introduction

Earlier chapters in this book deal with various aspects of the testing/evaluation/instruction linkage.

Chapter 2 describes a district in a crisis situation erupting from the public's reaction to publication of test scores. A number of issues are explored -- for example, the accuracy of the press' interpretation of the scores, the extent to which the scores serve as legitimate surrogates for student learning and as outcomes of the district's instructional effectiveness. More importantly, the scenario implies that the specific crisis is an instance of a more chronic problem -- one which is generic to many districts -- that is, the lack of a central district level organizational structure which routinely and effectively feeds back the evaluative data periodically collected to assess student achievement into the instructional system which impacts student achievement.

In Chapter 3, single factor solution strategies, often proposed by the public and even by some educators, as responses to the immediate problem are presented, discussed and rejected. We reject them because we believe that taken individually, each points towards a dead-end direction. That is, each of them may be an important element in an integrated data-based instructional management system. Singly, however, they are energy consuming diversions not making much of a contribution to educational environments which enhance student learning.

In Chapter 4, six districts who recognized the interaction between assessment and action, between testing and instruction, tell how they developed their own data-based management systems. One district's system looks very different from another district's system. In all cases, however, improved student learning was the aim of the system. The teacher was recognized as the key to the delivery of effective instruction. The role of the school as the setting within which teachers work was asknowledged. District responsibility for supporting the principal and the teacher with relevant information and with data-informed policies was asserted.

In Chapter 5, the generic difficulties experienced by school districts when they developed data-based instructional management systems are described. Many school districts these days exist in turbulent social, political and economic environments. Their budgets may decrease while their student needs increase. Their population may shift within the district causing some schools to close, others to go on double session. The first language of many students may not be English. Student stress may be acted out in terms of drug use, vandalism, absenteeism, violence. Such turmoil and the rapidity with which non-educational problems seem to accumulate may engender cynicism or despair on the part of school officials that they can educate students at all. For these educators, district management is focused on instructional survival rather than instructional improvement.

In our view, school districts' need to respond to rapidly changing environmental conditions is fuel for the development of a central office system for collecting, analyzing, and using test or other evaluative data to enhance student progress and factors that seemingly contribute to or impede that progress.

In this Chapter, then, we discuss specific ways in which districts who are interested in such a point of view can engage in action planning to create such a management system for themselves.

Definition of Action Planning

As we use it here, action planning is not the pre-ordinate specification of goals and objectives accompanied by a blueprint of activities and by a time/task line, although it might well have some of these characteristics. Action planning is not the arm-chair product of a single administrator's conception of a data-based instructional management system, although key individuals might well have decisive influence. On the other hand, neither is action planning a form of "muddling through" or "making it up as we go along" although at particular times and for particular individuals, it may well feel that way.

Action planning, as we infer it from the "heroic districts" who have described their efforts in Chapter 4, requires district leaders with the vision to so shape action and events that they cumulate into new structures. For example, in Los Alamitos, the unacceptability of the criterion-referenced tests first purchased by the district did nct cause the abandonment of the data-based instructional plan, but rather refocused action on developing a set of district objectives. In San Juan, the difficulties with Title I evaluations and program planning led not to abandonment of the idea of planning and evalution but to greater receptivity to the school site planning orientation of the California Early Childhood Education Program. In these two districts, as well as in the others in Chapter 4, individuals did not know at the outset of their endeavors precisely what their data mangement system would look like. Actions led to plans. These plans, in turn, gave rise to anticipated as well as unanticipated ideas and events which caused adjustments in subsequent actions and plans.

Action planning, therefore, can be characterized as simultaneously reactive and proactive, simultaneously spontaneous and controlled, combining or alternating both the doing and the reflecting. Action planning is a process which enables a district staff to accommodate to the complexities of the external environments with which they are in constant interaction. It seems consonant with the loose coupling of a school district's internal subsystems characterizing most school districts today.

Attitudes on Which District Action Planning Rest

The professional attitudes which often accompany district action planning to create data-based instructional management systems can be divided into four categories.

Attitudes towards data. District personnel must believe -- in whole or in part -- that

- data from tests and from other evaluative devices such as surveys, observations, demographic statistics, are tools for describing the current status of students, for analyzing that status and for deciding how to upgrade it
- data from tests and other evaluative devices can be collected, analyzed and disseminated so as to be of use to a variety of audiences -- e.g., teachers, principals, staff developers, curriculum specialists, administrators -- for a variety of purposes -- e.g., clinical, programmatic, managerial, policy
- the management of data from tests and other evaluative devices must be integrated with the management of classroom instruction, so that the two processes are compatible and interact with one another

Attitudes towards instruction. District personnel must belive -- in whole or in part -- that

- instruction can affect student learning in ways that can be measured by tests or other evaluative devices
- teachers can manage the elements of instruction -- e.g., materials, time on task, teaching strategy -- so as to affect student performance on tests

- teachers who deliver instruction to students are themselves part of and responsive to the school and the district context which influences how and what they teach
- the district central office has an important role -- facilitative or directive -- in the instructional process as it occurs in their schools; the district has overall responsibility for the learning of the students who attend its schools

Attitudes towards management. District personnel must believe -- in whole or in part -- that

- strong central office management -- of internal operations and of relationships to schools -- can upgrade teaching and therefore effect student learning
- central office management of tests, evaluations, curriculum, staff development, materials, can be integrated and coordinated through informal as well as formal mechanisms
- there is no single "right way" to develop a data-based instructional management system that can be imported from elsewhere and implanted within the district

Attitudes towards change. District personnel must believe -- in whole or in part -- that

- the development of a data-based instructional management system may require changes in behaviors or in attitudes on the part of administrators, principals, teachers, students
- change in school districts is possible. It requires careful attention to the factors and forces which have "frozen" the organization into its present configuration as well as understanding of the "unfreezing" and "refreezing" processes essential to organizational change and renewal
- it is likely that change in school districts will be incremental rather than radical, gradual rather than sudden, uneven rather than orderly
- change may produce failures as well as successes
- change may be threatening to some individuals or groups, energizing to others

Conditions Impeding or Facilitating Action Planning

There are many conditions within school district organizations which might impede the action planning needed to create data-based instructional management systems. We can list these briefly because they have already been discussed elsewhere.

- rapid changes make it difficult to undertake a coordinated long term effort to create a data-based instructional management system

- boundary permeability requires school districts to attend to or comply with legislative and administrative regulations whether or not these actually contribute to local educational improvement

- weak technical core for education. The fact that there is a very small number of research validated cause and effect relationships generalizable to all teaching-learning situations encourages teacher, principal and administrator reliance on their own working knoweldge and insistance on their right to do so.

- loose administrative coupling. Policy and administrative decisions taken at the board and central office levels may not filter down through school and classroom layers to directly affect students. In part, this may be due to imperfect communication devices, in part to differing role-related perspectives about what is important to do, in part to traditions within the American public education system.

There are, however, many conditions which facilitate action planning to link testing and evaluation with instruction. These can be devided into conditions external to the district and those internal to it. They will be presented here in the form of checklists.

External Conditions Favorable to Developing a Data-based Instructional Improvement System

☐ Stable student population with few sudden shifts in size or distribution

☐ Community support for public education

☐ Community agreement on desired educational outcomes for their children

☐ Community agreement on instructional methods

☐ Community acceptance of test scores as indicators of student learning

☐ Community willingness to support district policies

☐ Parent willingness to support instructional change

☐ Media encouragement and willingness to learn about test score/school practices

☐ Availability of external funds from governments, foundations, businesses, etc.

☐ Availability of consultant assistance from universities, professional organizations, county organizations, other districts

☐ Availability of research and development reports, journals, etc.

Internal Conditions Favorable to Developing a Data-based Instructional Improvement System

☐ Strong "idea champions" in positions of power or influence

☐ Stable core of supportive staff in central office and in schools

☐ Individuals with technical and computer skills related to construction, selection, analysis of test or other evaluative data

☐ Individuals with substantive and technical skills related to curricular development

☐ Individuals with management skills related to team building, consensus development, communications, etc.

☐ Teachers interested in instructional improvement willing to make changes in their own methods, materials, classroom management, etc.

☐ Financial resources within the district to pay for system development and maintenance

☐ Availability of computer services

Section 2

Alternative Action Plans
Discussion and Outlines

Discussion

Each of the four plans presented in outline form here represents a conglomerate of actions undertaken by districts in pursuit of their intended purpose.

As is evident from the stories presented in Chapter 4, not all districts start from a clearly defined, explicit purpose. Some districts act their way into new kinds of thinking, while other districts may think their way into new kinds of acting. However, it is usually possible to infer from the statements of key individuals, or from an examination of the activities, whether district intentions are narrowly defined in terms of raising student test scores or more broadly defined in terms of altering district management processes.

Often, however, district intentions evolve over time. If impetus for change comes, as we see in our Chapter 2 Scenario, from public outcry around test scores, the immediate response of the district may be test-focused. However, as individuals in the central office, schools and classrooms think about strengthening the interaction between what teachers do, what students learn, and the demonstration of learning on tests, district actions may move in new directions. This is the essence of action planning as we have defined it.

The four outlined plans are not exhaustive of all possible district foci, but they do seem characteristic of the range of school district action strategies.

The key difference among the plans is in the pivot point around which changes occur. "Possible payoffs" may appear similar among plans but the probability of their being achieved does differ from plan to plan. In Plan

A, for example, gains in student achievement test scores are highly likely in the short term whether or not important changes in instruction take place. By contrast, in Plan B, curriculum-based instructional changes may occur well in advance of any rise in student achievement scores. In Plan C, school site planning may well have beneficial consequences on school climate, teacher morale and student learning but these consequences may or may not be demonstrated by test score increases. The same is true, at the district level, for Plan D.

The bare-bones "action sequence" described in each plan are not prescriptive. They only indicate some of the steps that many districts have taken in pursuit of their explicit or implied purposes. Specific "how-to" suggestions which flesh out each of the plans are included in Section 3's Tips and Techniques.

Our discussion of "district responsibility," likewise, is suggestive of methods in use by districts. Each district is unique in terms of its management style, its past history, its current situation; so each district must work out its own arrangements relating to directing, facilitating, consulting with, informing, supervising, supporting various groups.

Finally, the "advantages" and "difficulties" are summary compilations of experiences which may be helpful.

Plan A: Test-Focused

Explicit Purpose - to raise test scores for all students or particular subgroup(s) (e.g., language, SES, grade level, school) in all or a subset of tested skills (e.g., reading, math, language arts)

1. Pivot point - currently administered test(s) (e.g., a norm-referenced test such as the CAT, MAT, CTBS, Iowa Test of Basic Skills; a locally developed or purchased criterion-referenced test battery; State assessment tests; district or State proficiency tests)
2. Possible payoffs - 2.0 Improved test scores
    - 2.1 Capability for tracking progress of individuals, subgroups, or total population over time; tracking of teachers, grades, schools by student scores
    - 2.2 Provision of information to parents, board, media about student achievement as demonstrated by test scores
    - 2.3 Improved instruction
3. Action sequence - 3.0 Identify test or subtest(s) on which scores are to be raised
    - 3.1 Describe target students whose scores are to be raised
    - 3.2 Analyze knowledge or skills assessed by test
    - 3.3 Analyze past/current scores of students
    - 3.4 Act: provide students with practice in test-taking skills, and/or provide students with practice on items similar to those on the test, and/or provide students, prior to test, with instruction and materials targeted to knowledge and skills tested
    - 3.5 Analyze test score results, disseminate and plan for additional action
4. District responsibilities

    District administrators assume responsibility for raising student scores. With appropriate input from curriculum and testing specialists, from principals and teachers, they identify the test and the students. Analysis of the tested knowledge and skills can be done in the district office or by committees of teachers. Analysis of past and current scores might be handled by a specialist. Development of action strategies (and the supporting materials or procedures to implement them) can be handled either centrally, or by school, or by teacher committee.
5. Advantages - 5.0 Short term targeted practice for students on tested skills is likely to raise their scores .... expensively, without much system change.

6. Difficulties - 6.0 Emphasis on raising test scores may diminish attention to non-tested learnings.

6.1 Emphasis on raising test scores may encourage cheating; cause unproductive anxiety over test performance in teachers and children; cause excessive reliance on single indicator of student abilities.

Plan B: Curriculum/Instruction-Focused

Explicit Purpose - to improve classroom instruction for all students or particular subgroup(s) (e.g., by language, SES, grade level, school), for all or a subset of specified objectives

1. Pivot point - current teaching practices and materials relevant to specified objectives

2. Possible payoffs -
   - 2.0 Improved instruction on particular objectives
   - 2.1 Improved student achievement scores on currently administered test(s)
   - 2.2 Upgraded teacher skills
   - 2.3 Improved classroom management
   - 2.4 Improved classroom climate
   - 2.5 Improved support system for teachers
   - 2.6 Improved supervision system for teachers

3. Action sequence -
   - 3.0 Describe target student group(s)
   - 3.1 Describe desired student outcomes, e.g., in reading, math, language arts
   - 3.2 Select or develop tests, evaluative outcome measures (e.g., CRT's, text-tests, teacher tests, surveys, unobtrusive measures such as rates of absenteeism, vandalism, etc.)
   - 3.3 Analyze materials/instruction in relation to desired student outcomes
   - 3.4 Describe base line
   - 3.5 Act: coordinate texts:tests:teaching, and/or modify teaching through training/supervision, and/or provide supports to teachers: aides, pe....sts, materials, etc.
   - 3.7 Evaluate using outcome measures and disseminate results
   - 3.8 Take additional action

4. District responsibilities

   District administrators assume responsibility for improving instruction. With appropriate input from curriculum, testing, evalution, staff development specialists, principals and teachers, they identify curricular objectives, present instruction (teaching methods/materials) as well as the changes in instruction which are called for. Central office staff coordinate their own operations in areas such as curriculum, testing, supervision, staff development in order to monitor, train, support, and assist teachers.

5. Advantages -
   - 5.0 Attention is focused on improving those classroom instructional practices including teaching methods, time-on-task, materials, etc., that are relevant to particular objectives, particular groups of students.

6. Difficulties - 6.0 Test scores for norm-referenced tests may remain low if such tests are mismatched with district curricular objectives or text materials.
6.1 Instructional focus requires teacher support and willingness to adopt new methods, materials, etc.
6.2 District-level coordination of support and supervisory relationships with principals and teachers may be difficult.

Plan C: School-Focused

Explicit Purpose - to improve school site instructional planning and assessment

1. Pivot point - current school management

2. Possibile payoffs
   - 2.0 Improved school management
   - 2.1 Involvement of principal, teachers, community in teams to develop school-wide problem identification, problem resolution, resource allocation
   - 2.2 School programming adapted to local community needs
   - 2.3 Improved student achievement scores on currently administered tests
   - 2.4 Improved student knowledge, skills, attitudes on other measures
   - 2.5 Improved support system for teachers
   - 2.6 Improved supervision system for teachers

3. Action sequence
   - 3.0 Describe target schools
   - 3.1 Do needs assessments/problem surveys
   - 3.2 Identify high priority problems
   - 3.3 Analyze existing data about school-level performance on high priority areas or problems
   - 3.4 Describe desired outcomes: student, class-level, school-wide
   - 3.5 Select or develop tests or other evaluative outcome measures
   - 3.6 Identify and choose among alternative actions
   - 3.7 Act
   - 3.8 Evaluate using outcome measures
   - 3.9 Analyze and disseminate results
   - 3.10 Take additional action

4. District responsibilities

   District administrators assume responsibility for improving school site planning and assessment. With appropriate input from community, principals, teachers, parents, district facilitates the organization of school site teams, provides them with the needed training, technical support, resources to do their own problem identification, action planning, implementation and assessment.

5. Advantages
   - 5.0 School effectiveness may be increased through community support, regular planning and assessment procedures, ownership of solutions by principal and teachers, etc.
   - 5.1 Schools can develop solutions responsive to their own situation.
   - 5.2 Process may lead to increased morale, feelings of efficacy, school spirit, etc.

6. Difficulties - 6.0 Student test score increases may or may not be a high priority school goal; test scores may or may not be affected -- short term or long term -- by school site plans.

6.1 Leadership, composition, productivity of school planning groups may vary greatly.

6.2 Support, technical assistance, resources available to schools may be costly compared to observable benefits.

6.3 Gains may be long term rather than short term.

6.4 Sustaining time-intensive efforts may be difficult.

Plan D: District Management-Focused

Explicit Purpose - to coordinate all district operations towards the enhancement of student learning

1. Pivot point - current district operations (e.g., testing, evaluation, curriculum, instruction, materials and media, supervision, staff development

2. Possibile payoffs
   - 2.0 Improved district management
   - 2.1 Improved accountability system for principals, teachers
   - 2.2 Coordination of testing, evaluation, curriculum, instruction, texts, etc.
   - 2.3 Upgraded teaching skills
   - 2.4 Upgraded principals management skills
   - 2.5 Improved public image
   - 2.6 Improved district efficiency
   - 2.7 Improved student achievement scores on existing tests

3. Action sequence
   - 3.0 Identify and publicize district philosophy in terms of student learning, teaching processes, management characteristics
   - 3.1 Examine current situation in one or more areas by analyzing existing data, collecting additional information from needs assessments, problem surveys, instructional analyses, test scores, etc.
   - 3.2 Develop outcome measures, other indicators of progress
   - 3.3 Develop action plan for policy, administrative, school and classroom levels
   - 3.4 Act
   - 3.5 Monitor progress using outcome measures and indicators, and reassess
   - 3.6 Take additional action

4. District responsibilities - District administrators assume responsibility for coordinating internal operations as well as school relationships in such a way that student learning is enhanced. District takes leadership in identifying the needs, problems, associated with improving student learning; in directing or facilitating coordinated action planning to meet them; in assessing progress.

5. Advantages - 5.0 Examination and redirection of district management will lead to more effective central office leadership, supervision, information collection and analysis, etc.

6. Difficulties - 6.0 Test scores may not be affected in the short term by more efficient district management.

Section 3

Tips and Techniques: Criterion-referenced and Norm-referenced Tests

Criterion-referenced Tests

How To Decide If You Should Consider a Criterion-referenced Testing Program

(If most of your answers are unboxed, you probably don't want a CRT system. If most of your answers are in the boxes, you should assume that a CRT system may be useful for your district.)

1. Do the tests that you now give -- either district-wide or by teacher choice -- provide the teacher with information that helps him/her tailor instruction to student needs? $\boxed{\text{NO}}$ YES

2. Do the tests that you now give -- either district-wide or by teacher choice -- provide the teacher with information that helps him/her communicate with parents, aides, other teachers, about student needs? $\boxed{\text{NO}}$ YES

3. Do the tests that you now give -- either district-wide or by teacher choice -- provide the principal, teacher or parent with information that helps him/her understand individual student progress across grade levels? $\boxed{\text{NO}}$ YES

4. Do the tests that you now give -- either district-wide or by teacher choice -- provide the principal with information about the learning of students by class, or by grade level, or by some other relevant subgroup? $\boxed{\text{NO}}$ YES

5. Do you have, or want to develop, a minimum or standard set of objectives that all, or identified subsets of, students will achieve? NO $\boxed{\text{YES}}$

6. Do you have, or want to develop, a cross-referenced system that makes tests, texts, teaching time consistent with a set of district-wide objectives? NO $\boxed{\text{YES}}$

7. Do you have, or want to develop, a set of remediation techniques -- e.g., materials, programs, teaching methods, -- for students who are non-masters of particular objectives or groups of objectives? NO $\boxed{\text{YES}}$

8. Do you have, or want to develop, positive teacher attitudes and teacher skill in "teach-test-reteach" and "diagnostic/prescriptive" teaching strategies? NO $\boxed{\text{YES}}$

Criterion-referenced Tests

How To Develop a Criterion-referenced Testing System

1. Recognize that the development of such a system takes time -- many districts have spent more than five years in developing a system.

2. Recognize that there are tradeoffs for any major decision, and that a district's best judgment combined with knowledge of the benefits and costs is the best available method for making the decision.

3. Allocate roles and responsibilities for developing the system.

4. Keep all relevant groups -- Board, unions, press, teachers, parents, etc., informed as appropriate.

5. Decide on a start-up development strategy, either 1) one subject (e.g., math, reading, language arts) for all grades; 2) major subjects for one grade.

6. Compile a district curriculum scope and sequence by subject matter and grade level, stated in terms of student outcomes, by any of the following methods:

    a. use available subject matter scope and sequence done by text publisher;

    b. borrow another district's scope and sequence and adopt or modify;

    c. infer from teachers what is already being done by subject by grade, write it down and refine it in terms of student outcomes;

    d. hire outside consultants to create sequence;

    e. train teacher committees to write student outcomes by grade and subject area.

7. Identify important end-of-semester learning objectives, and develop item clusters for each objective.

8. Identify important en-route objectives and develop item clusters for each objective.

9. Have teachers/subject matter specialists view items or revise.

10. Try out items with students and revise.

11. Package items into test booklets.

12. Develop administration procedures, e.g., timing and conditions of testing, and orientation for teachers.

13. Develop scoring, and feedback procedures, e.g., turnaround time, format for reviewing student, subgroup, class, grade achievement.

14. Develop remediation policies for non-mastery students at group, class, grade, school, district level, as appropriate.

15. Develop procedure to identify and handle objectives and tasks that need revision.

Criterion-referenced Tests

How To Purchase a Criterion-referenced Test System

1. Locate the objectives that the test items have been written to measure.
2. Determine whether the tested objectives master district objectives; if not, is the district willing to change instruction to match test's objectives?
3. Determine whether written items match stated objectives, i.e., can you infer objectives from items?
4. Determine whether objectives and the related test items are clustered in the sequence in which your district teaches the objectives.
5. Examine the items. (Is the level of vocabulary, sentence structure, format, appropriate to the grade level and students you want tested?)
6. Determine whether test administration instructions are clear.
7. Determine whether the scoring of student answers to items is quick, easy, easy for teachers to interpret.
8. Determine whether there are alternate items available for pre/post or retest purposes.
9. Ask teachers to examine specimen sets of items. (What problems do they see? Can the problems be resolved?)
10. Find out about the services provided by the test company -- training, scoring, interpretation, development of new items, tailoring of tests to your needs, etc. -- and their costs.
11. Weigh costs of purchasing tests against costs of developing tests. (Think about dollars for salaries and time, available personnel, level of commitment, tradeoffs with other activities, length of time till payoff, etc.)

Criterion-referenced Tests

How To Construct Your Own Criterion-referenced Tests

1. Specify the objectives the district wants tested. (These may be end-of-semester and/or en-route objectives.)

2. Describe the specifications for the items related to each objective -- that is, their content, the vocabulary. the types of distractors (for multiple choice formats), the criteria used in grading (for essay or short answer formats), etc.

3. Develop item formats and sample items that describe a) instructions to students; b) stimula limits; and c) response limits.

4. Write several items for each objective.

5. Review items for accuracy of content and instructional ambiguities as well as their conformity to specifications.

6. Mix up the items and ask teachers to match items to objectives, or to infer objective from item. Eliminate items that seem irrelevant to objective.

7. Ask students to answer all items pertaining to one objective. Review and revise items that seem to cause trouble, paying particular attention to those items missed by students who answered most items correctly.

8. Remember that you want a group of items, all of which measure the same objective, that students who have mastered the objective will answer correctly, and that students who have not yet mastered the objective will answer incorrectly. Ideally, you would like some multiple-choice items so that student errors will provide you with clues what about the non-mastery student has misunderstood or missed.

How To Determine Test Length

1. Recognize that test length is a tradeoff between getting an accurate measure of students' knowledge and skills (the more items, the more reliable the test), and other uses for the time spent on testing.

2. Three or four items per objective seems to be practical and feasible for most situations. More items should be used when important decisions are to be based on student test performance.

How To Determine Passing Scores

1. Recognize that passing scores are arbitrary and that mistakes can be made.

2. Set what appear to be reasonable cut-off points, e.g., 3 out of 4 items; 80 percent correct. Examine whether those students who "clearly pass" and those who "clearly fail" would fall into the same categories if measured with some other criterion such as teacher judgment, classroom or homework performance, etc. Readjust cut-off point or revise items that do not seem to discriminate between "clear passes" and "clear failures."

3. Set up review procedures for those students who fall between those who "clearly pass" and those who "clearly fail."

4. Consider what remedies will be made available to non-passing students.

Criterion-referenced Tests

How To Support an On-going Criterion-referenced Testing System

1. Recognize that the tests themselves, keyed to a consensus set of objectives, are only one part of improving instruction through management of testing and evalution.

2. Key curricular objectives and the criterion-referenced tests to materials/media.

3. Key curricular objectives and the criterion-referenced tests to instructional time and methods.

4. Develop procedures to find and handle difficulties with items, formats, test administration, feedback to interested audiences.

5. Sensitize groups to ways in which test data analyses can inform policy making and administration in areas such as staff development, budgeting, school planning, communication to public and media.

6. Remain aware of possible undesirable side effects, among them

    a. teaching the test items

    b. narrowing instruction to what is tested

    c. neglecting high achieving students

    d. devoting undue time to testing

    e. fragmenting instruction into measurable pieces

7. Build in periodic review of curricular objectives and tests.

Norm-referenced Tests

How To Decide If You Should Consider a Norm-referenced Testing Program

If most of your answers are unboxed, you probably don't want an NRT program. If most of your answers are in the boxes, you should assume that an NRT system may be useful for your district.)

1.	Do the tests that you now give -- either district-wide or by teacher choice -- provide the teacher with information that helps him/her tailor instruction to student needs?	$\boxed{\text{NO}}$	YES
2.	Do the tests that you now give -- either district-wide or by teacher choice -- provide the teacher with information that helps him/her communicate with parents, aides, other teachers, about student needs?	$\boxed{\text{NO}}$	YES
3.	Do the tests that you now give -- either district-wide or by teacher choice -- provide the principal, teacher or parent with information that helps him/her understand individual student progress across grade levels?	$\boxed{\text{NO}}$	YES
4.	Do the tests that you now give -- either district-wide or by teacher choice -- provide the principal with information about the learning of students by class, or by grade level, or by some other relevant subgroup?	$\boxed{\text{NO}}$	YES
5.	Do you have, or want to develop, a minimum or standard set of objectives that all, or identified subsets of, students will achieve?	$\boxed{\text{NO}}$	YES
6.	Do you have, or want to develop, a cross-referenced system that makes tests, texts, teaching time consistent with a set of district-wide objectives?	$\boxed{\text{NO}}$	YES
7.	Do you have, or want to develop, a set of remediation techniques -- e.g., materials, programs, teaching methods, -- for students who are non-masters of particular objectives or groups of objectives?	$\boxed{\text{NO}}$	YES
8.	Do you have, or want to develop, positive teacher attitudes and teacher skill in "teach-test-reteach" and "diagnostic/prescriptive" teaching strategies?	$\boxed{\text{NO}}$	YES

9. Do you need a test that will compare district students' performance with others of similar background on particular skill areas?	NO	YES
10 Is a norm-referenced test required by law or regulation to evaluate federal or state programs?	NO	YES
11. Do the Board, the public, the media, parents, colleges, important others, insist on comparative data on a nationally-known test?	NO	YES
12. Can you locate a norm-referenced test which meets the minimum requirements of validity for your district's student population, curricular objectives?	NO	YES

Norm-referenced Tests

How To Purchase a Norm-referenced Test System

1. Choose a test that is matched as closely as possible to district curriculum objectives. To determine this:
   a. classify the test items from the test form according to district curricular objectives;
   b. tally the number of district objectives for which there are one, several, or many test items included the test;
   c. tally the number of test items connected to those district objectives taught to students prior to testing.

2. Determine from the manual if the test is normed with students similar to those in your district at the time of year you plan to give the test. (If not, consider how this would affect your interpretion of the scores, if you use this test.)

3. Review the test for negative factors such as
   a. test bias in content, vocabulary, pictures, etc.
   b. confusing test administration requirements
   c. confusing test format/instructions to students
   d. higher than usual per-pupil costs (20 cents per pupil is low, 80 cents at the high end)

4. Find out what services and score analyses are available from publishers.

5. Involve a variety of individuals in test selection procedures, e.g., curriculum and test specialists, administrators, teachers, parents, etc.

6. Review as many relevant tests as possible. Get names from Buro's Mental Measurements Yearbook, published by Gryphon Press, Highland Park, New Jersey.

Norm-referenced Tests

How To Interpret Norm-referenced Test Scores

1. A student's percentile score includes the percentage of the norm group whose raw scores fall below the student's raw score. For example, performing at the 60th percentile means that a student's raw score was higher than 60 percent of the students in the norm group. Test publishers should provide percentile scores for grades or schools as they cannot be obtained by averaging individual percentiles.

2. A student's standard score indicates how far above or below the norm group mean of raw scores that student's raw score fell.

3. A student's stanine score also indicates how far a student's raw score deviates from the norm group mean of raw scores. The raw scores are divided into nine intervals containing a fixed percentage of the raw scores. For example, 4 percent of the raw scores fall in stanine 1 and in 9; 7 percent of the raw scores fall in stanine 2 and in 8; 12 percent in stanine 3 and in 7; 17 percent in stanine 4 and in 6; 20 percent in stanine 5. If a student has a stanine of 9, he/she has done better than 96 percent of normed students.

4. A student's grade or age equivalent score tells where his/her raw score falls with respect to the average performance of students at various grades or age levels. There are many methodological problems with these scores and many experts recommend they not be used.

5. Consult with test publishers to determine most useful analyses for district purposes, e.g., by individual, class, grade, school; by student characteristics of instructional relevance, such as language, SES, length of time in school, etc.; by subtests.

Norm-referenced Tests

How To Use Norm-referenced Test Scores

1. To group or place students. Examine scores on particular content or subject areas and group students with similar skill profiles or group heterogeneously according to some instructional philosophy.

2. To diagnose student needs. Examine subtest results to determine individual pattern of achievement/deficiency. Look for surprises, that is, students who do better or worse than expected on all or some subtests.

3. To evaluate instructional programs. Compare actual performance of students with expected performance. Examine reasons for correspondence discrepancy, e.g., is it in the test? the instructional program? or the match between test and instructional program?

   Or, compare students' gains from year to year to get gross estimates of actual performance versus expected performance over time. Examine reasons for correspondence or for discrepancy.

Tips and Techniques: Evaluations

Evaluations (List of topics)

- How to develop non-test evaluations of student achievement
- How to develop non-test evaluations of instructional programs
- How to regularize collection of non-test evaluative data
- How to disseminate findings from non-test evaluative data
- How to insert evaluative data into policy and administrative decision making

Tips and Techniques: Curriculum and Instruction

Curriculum and Instruction (List of topics)

- How to decide if you need a uniform set of curricular objectives
- How to develop curricular objectives by subject areas, by grade levels
- How to identify milestone objectives for which test items could be developed
- How to index curricular objectives to text and other materials
- How to assess instructional time/amount of practice students receive on curricular objectives

Tips and Techniques: School Site Planning and Assessment

School Site Planning and Assessment: (List of topics)

- How to decide on the desirability of school site planning and assessment
- How to initiate school site planning and assessment
- How to form, facilitate, train and support school site teams
- How to determine data needs of and data presentations to school site teams
- How to assess whether school site teams are cost effective

Tips and Techniques: District Management

District Management: (List of topics)

- How to identify needs for coordination among district operations such as testing, evaluation, curriculum, instruction, staff development
- How to select and implement appropriate coordinating mechanisms among district operations
- How to collect, analyze and disseminate data useful in policy decisions
- How to collect, analyze and dissiminate data useful in school improvement
- How to collect, analyze and disseminate data useful in classroom decision making
- How to provide school-level supports, supervision and training so that principals and teachers use test and evaluative data to improve instruction

Section 4

Initiating Action Planning: Discussion and Worksheets

Discussion

The six worksheets here are intended as guides or agendas for groups of people assembled by district administrators for the purpose of initiating planning to improve instruction through the management of testing and evaluation activities.

It is anticipated that many of the conditions facilitating action planning are present in the district before this process is begun. For example:

1) an external environment which is supportive of instructional improvement, and the use of test scores and other evaluative data to report on student progress

2) idea champions and supporters within the district who have the motivation, skills and power to move the district in this direction

3) resources and time to consider alternative action plans and to manage the implementation of decisions resulting from that consideration

Worksheet #1 examines current practices. Its intended outcome is based on understanding of constraints -- those that are fixed and those that are alterable.

Worksheet #2 explores satisfaction and dissatisfaction of various individuals and groups with current practice. Its intended outcome is identification of agreed-upon troublesome areas or problems.

Worksheet #3 surfaces recent attempts or thoughts about change as well as the supporting or opposing parties. Its intended outcome is arraying the reasons why changes have or have not come about.

Worksheet #4 elicits ideas and suggestions for action planning and identifies where support or resistance is likely to come from. Its intended outcome is an informal identification of priorities.

Worksheet #5 is a force field analysis for high priority areas. Its intended outcome is a detailed description of factors influencing action planning.

Worksheet #6 is a further listing of factors influencing action planning. Its intended outcome is the formation of an action plan.

1. WHAT IS THE DISTRICT NOW DOING IN THE AREAS OF TESTING, EVALUATION, CURRICULUM, INSTRUCTION?

1. What *tests* is the district now using? For each of the tests, think about: 1) who/what influenced the selection of that test; 2) who is now interested in having the test given; 3) who now sees the test scores; 4) what use, if any, do various groups of people (e.g., parents, teachers, principals, central office, press) make of the scores; 5) what would happen if the test were *not* given?

2. What *evaluations* does the district now do? For each evaluation activity, think about: 1) why is the evaluation done? 2) Who is interested in having the evaluation done? 3) What use, if any, do various groups of people (e.g., parents, teachers, principals, central office, press) make of the findings? 4) What would happen if the evaluations were *not* done?

3. What *curriculum* does the district now have? What are its characteristics (e.g., to what extent does it guide teachers' choice of content, skills; to what extent is it consistent with texts, tests; to what extent is it differentiated for subgroups of students)? What ould happen if the test were not given?

4. What *instructional* policies does the district now have? To what extent are they implemented in the classroom? Who/What influenced the development of the policies? Who oversees their implementation? What would happen if the instructional policies were changed?

ANALYSIS	What is now being done that must remain in place? What changes could be made? What factors must be considered when making changes?

2. HOW SATISFIED ARE VARIOUS GROUPS WITH WHAT THE DISTRICT IS NOW DOING IN TESTING, EVALUATION, CURRICULUM, INSTRUCTION?

TESTS

1a. What are areas of satisfaction with current tests?	Who holds these views?	1b. What are areas of dissatisfaction with current tests?	Who holds these views?
________	________	________	________
________	________	________	________
________	________	________	________

EVALUATION

2a. What are areas of satisfaction with current evaluation activities?	Who holds these views?	2b. What are areas of dissatisfaction with current evaluation activities?	Who holds these views?
________	________	________	________
________	________	________	________
________	________	________	________

CURRICULUM

3a. What are areas of satisfaction with current curriculum activities?	Who holds these views?	3b. What are areas of dissatisfaction with current curriculum activities?	Who holds these views?
________	________	________	________
________	________	________	________
________	________	________	________

INSTRUCTION

4a. What are areas of satisfaction with current instructional activity?	Who holds these views?	4b. What are areas of dissatisfaction with current instructional activity?	Who holds these views?
________	________	________	________
________	________	________	________
________	________	________	________

ANALYSIS	What are areas of consensus and development among various groups? What are major areas of satisfaction? Major areas of dissatisfaction?

3. WHAT CHANGES HAVE BEEN RECENTLY CONSIDERED OR MADE IN TESTING, EVALUATION, CURRICULUM, INSTRUCTION?

1. in the area of testing: (e.g., eliminating, replacing, adding tests; changing dates; changing analysis; changing dissemination of results, etc.)

changes considered	who suppported/why?	who opposed/why?	what happened?

2. in the area of evaluation: (e.g., eliminating, modifying, adding evaluations; changing questions, design, instruments, analysis, dissemination of results, etc.)

changes considered	who suppported/why?	who opposed/why?	what happened?

3 in the area of curriculum: (e.g., eliminating, revising, redoing)

changes considered	who suppported/why?	who opposed/why?	what happened?

4. in the area of instruction:

changes considered	who suppported/why?	who opposed/why?	what happened?

ANALYSIS	What changes have been made and what were the reasons they were made? Who were internal and external "idea champions"? Supporters? How did they influence the changes?

4. IN THE AREAS OF TESTING, EVALUATION, CURRICULUM, INSTRUCTION WHAT WOULD VARIOUS GROUPS LIKE TO HAPPEN?

The community/press would like	
The school Board would like	
Central Office staff would like e.g.__________ __________ __________	
Principals would like	
Teachers would like	
Students would like	
Parents would like	

ANALYSIS	What are areas of consensus and disagreement among various groups? Can desires be categorized: Easy/Difficult; Inexpensive/Expensive; Long-term/Short-term? What high priority areas for action planning can be identified?

5. FOR EACH HIGH PRIORITY AREA LIST FACTORS FACILITATING/OPPOSING CHANGE AND THEIR STRENGTHS

CHANGE

Strength			Factors Facilitating Change	Factors Opposing Change	Strength		
High	Med.	Low			High	Med.	Low

ANALYSIS	Can the factors facilitating change be increased; factors opposing change decreased? Among all the priority areas which seems the most feasible plan to begin?

6. FOR THIS HIGH PRIORITY AREA, SHOULD THE DISTRICT BEGIN ACTION PLANNING?

1. State the explicit purpose for district action
   Who shares them?
   Who opposes them?

2. State the auxiliary payoffs desired by the district
   Who wants them?
   Who opposes them?

3. What are cost estimates for action?

4. What needed resources, e.g., skilled personnel, materials, equipment, etc., are available?

5. What is a reasonable timeline for action?

6. What should the roles and responsibilities of the Board be, if any?

7. What should the roles and responsibilities of district administrators be, if any?

8. What should the roles and responsibilities of principals be, if any?

9. What should the roles and responsiblities of teachers be, if any?

10. What problems can be anticipated?

11. How will outcomes of the actions be assessed?

ANALYSIS	Is there sufficient clarity of purpose, support, resources role definition to move on this priority?

REFERENCES

Alkin, M.C., Daillak, R. and White, P. Using evaluations: does evaluation make a difference? Beverly Hills, CA: Sage Publication, 1979.

Anderson, J.G. Bureaucracy in education. Baltimore: Johns Hopkins Press, 1968.

Bank, A. and Williams, R. Evaluation in school districts: organizational perspectives. Monograph No. 10, Los Angeles, CA: UCLA Center for the Study of Evaluation, 1981.

Bentzen, M.M., et al. Changing schools: the magic feather principle. New York: McGraw-Hill, 1974.

Berman, P. Educational change: an implementation paradigm. In R. Lehming and M. Kane (Eds.) Improving schools using what we know. Beverly Hills, CA: Sage Publications, pp. 253-286, 1981.

Berman, P. and McLaughlin, M.W. Federal programs supporting educational change, Vol. VIII: implementing and sustaining innovations. R-1589/8-HEW. Santa Monica, CA: Rand, 1978.

Berman, P. and McLaughlin, M.W. An exploratory study of school district adaption. R-2010 NIE. Santa Monica, CA: Rand, 1979.

Bidwell, C.E. The school as a formal organization. Chapter 23 in Handbook of organizations. Chicago: Rand McNally & Co., 1965.

Boruch, Robert F. and Cordray, David S. Report on use of evaluations at federal, state and local levels. September, 1980, (unpublished).

Boyd, W.L. The changing politics of curriculum policy-making for American schools. Review of Educational Research, 48 (4): 577-628, 1978.

Burns, T. and Stalker, G.M. The management of innovation. London: Tavistock Publications, 1961.

Carlson, R.O. Environmental constraints and organizational consequences: the public school and its clients. In D.E. Griffiths (Ed.) Behavioral science and educational administration. 63rd Yearbook, National Society for the Study of Education. Chicago: University of Chicago Press, pp. 262-276, 1964.

Clark, D. In consideration of goal free planning: the failure of traditional planning systems in education. In D. Clark, S. McKibbin and M. Malkar (Eds.) New Perspectives on Planning in Educational Organizations. San Francisco, CA: Far West Regional Laboratory, pp. 1-17, 1980.

Reference (continued)

Cohen, M., March, J. and Olsen, J. A garbage can model of organizational choice. Administrative Science Quarterly, 17: 1-25, 1972.

Corwin, R.G. The school as a formal organization. In L.C. Deighton (Ed.) Encyclopedia of Education. New York: Macmillan and Free Press, 8, 1971.

Corwin, R.G. Models of educational organization. Review of research in education. New York: American Educational Research Association, 7: 247-248, 1974.

Daft, R.L. and Becker, S.W. The innovative organization: innovation adoption in school organizations. New York: Elsevier, 1978.

David, J. Local uses of Title I evaluation. Palo Also, CA: Stanford Educational Policy Research Center, 1978.

Deal, T.E. Linkage and information use in educational organizations. Paper read to American Educational Research Association meetings, 1979.

Downs, G.W. Jr. and Mohr, L.B. Conceptual issues in the study of innovation. Administrative Science Quarterly, 21: 700-714, 1976.

Elboim-Dror, R. Some characteristics of the education policy formation system. Policy Sciences, 1: 231-253, 1970.

Farrar, E., DeSanctis, J. and Cohen, D.K. Views from below: implementation research in education. Cambridge, MA: The Huron Institute, 1979.

Firestone, W.A. and Herriott, W.A. Images of the school: an exploration of the social organization of elementary, junior high and high schools. Philadelphia, PA: Research for Better Schools, 1980.

Goodlad, J. The dynamics of educational change: towards responsive schools. New York: McGraw-Hill, 1975.

Goodlad, J. What schools are for. Bloomington IN: Phi Delta Kappa Educational Foundation, 1979.

Goodlad, J. and Klein, M.F. Behind the classroom door. Belmont, CA: Wadsworth Publishing Co., 1970.

Gouldner, A. Patterns of industrial bureaucracy. New York: The Free Press, 1954.

House, E. Three perspectives in innovation: technological, political and cultural. In R. Lehming and M. Kane (Eds.) Improving schools using what we know. Beverly Hills, CA: Sage Publications, 1981.

Reference (continued)

Lawrence, P. and Lorsch, J. Organization and environment: management differentiation and integration. Cambridge, MA: Harvard University Graduate School of Business Administration, 1967.

Lindbloom, C.E. The science of muddling through. Public Administration Review, 19, Spring 1959.

Lortie, D.C. School teacher. Chicago, IL: University of Chicago Press, 1975.

Lyon, C.D., Doscher, L., McGranahan, P. and Williams, R. Evaluation and school districts. Los Angeles, CA: Center for the Study of Evaluation, UCLA Graduate School of Education, 1978.

March, J.G. Footnotes to organizational change. Project Report No. 80-A6, Institute for Research in Educational Finance and Governance, Stanford University, 1980.

McLaughlin, M.W. Evaluation and reform: ESEA Title I, 1965. A Rand Educational Policy Study. Cambridge, MA: Bollinger Publishing Co., 1975.

Meyer, J.W. Research on school and district organization. Paper presented at Annual Sociology of Education Association Conference, San Diego, CA, 1977.

Meyer, J.W. and Rowan, B. Institutionalized organizations: Formal structure as myth and ceremony. American Journal of Sociology, pp. 340-363, September 1977.

Meyer, J.W. and Rowan, B. The structure of educational organizations. In M.W. Meyer and Associates (Eds.) Environments and organizations. San Francisco, CA: Jossey-Bass, pp. 78-107, 1978.

Miles, M. Mapping the common properties of schools. In R. Lehming and M. Kane (Eds.) Improving schools using what we know. Beverly Hills, CA: Sage Publications, 1981.

O'Shea, D. School district evaluation units: problems and possibilities. In A. Bank and R. Williams (Eds.) Evaluation in school districts: organizational perspectives. Los Angeles, CA: UCLA Center for the Study of Evaluation, pp. 117-133, 1981.

Patton, M.Q. Utilization-focused evaluation. Beverly Hills, CA: Sage Publications, Inc., 1978.

Perrow, C. Complex organizations: a critical essay. Glenview, IL: Scott Foresman, 1979.

Reference (continued)

Pincus, J. Incentives for innovation in the public schools. Review of Educational Research, 44: 113-144, 1974.

Pincus, J. and Williams, R.C. Planned change in urban school districts. Phi Delta Kappa, pp. 729-733, June 1979.

Rogers, E.M. and Rogers-Agarwala, R.R. Communications in organizations. New York: The Free Press, 1976.

Sarason, S.B. The culture of the school and the problem of change. Boston, MA: Allyn and Bacon, 2nd Edition, 1982.

Sieber, S. Knowledge utilization in public education: incentives and disencentives. In R. Lehming and M. Kane (Eds.) Improving schools using what we know. Beverly Hills, CA: Sage Publications, 1981.

Tomblin, E. and Harding, E. Small school district utilization of program evaluation: a survey of board members and superintendents. Paper presented at the American Educational Research Association annual meeting, Boston, MA, April 1980.

Weick, K.E. Educational organizations as loosely coupled systems. Administrative Science Quarterly, 21: 1-19, 1976.

Wise, A.E. The hyper-rationalization of American education. New York University Education Quarterly, 8(4): 2-6, 1977.

Zucker, L.G. Institutional structure and organizational processes: the role of evaluation units in schools. In A. Bank and R. Williams (Eds.) Evaluation in school districts: organizational perspectives. Los Angeles, CA: UCLA Center for the Study of Evaluation, pp. 69-90, 1981.